

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 12:59:40 ; Search time 55 Seconds
(without alignments)
570.232 Million cell updates/sec

Title: US-09-784-553C-19
Perfect score: 605
Sequence: 1 KPGRTVNLQYLHKNVVKAL.....IVLMAQTLEKIFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

ALIGNMENTS			
RESULT 1			
AA07027			
ID	AA07027	standard; protein; 754 AA.	
XX	AA07027;		
AC	AC		
DT	02-JUL-1999	(first entry)	
XX			
DE	Breast cancer associated antigen precursor sequence.		
XX			
KW	Cancer associated antigen; diagnosis; research; treatment; human;		
KW	breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;		
KW	prostate cancer.		
XX			
OS	Homo sapiens.		
XX			
PN	WO9904265-A2.		
XX			
PD	28-JAN-1999.		
XX			
PF	15-JUL-1998;	98WO-US014679.	
XX			
PR	17-JUL-1997;	97US-00896164.	
PR	10-OCT-1997;	97US-0061599P.	
PR	10-OCT-1997;	97US-0061765P.	
PR	10-OCT-1997;	97US-00948705.	
PR	11-OCT-1997;	97GB-00021697.	
PR	22-JUN-1998;	98US-00102322.	
XX			
PA	(LUDW-) LUDWIG INST CANCER RES.		
XX			
PI	Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;		
PI	O'hare M, Obata Y, Pfeundscht M, Tureci O, Sahin U;		
XX			
DR	WPI; 1999-132448/11.		
XX			
PT	New isolated cancer associated nucleic acids and polypeptides - isolated		
PT	using sera from cancer patients, used to develop products for the		
PT	diagnosis, monitoring or treatment of cancers.		
XX			
PS	Disclosure; Page 404-405; 787pp; English.		
XX			
CC	The invention relates to a method for diagnosing a disorder characterised		
CC	by expression of a human cancer associated antigen precursor coded for by		
CC	a nucleic acid molecule (NAM). The method comprises: (a) contacting a		
CC	biological sample isolated from a subject with an agent that specifically		
CC	binds to the NAM, an expression product or a fragment of an expression		
CC	product complexed with an HLA molecule; and (b) determining the		

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	605	100.0	754 2 AA07027 Breast ca
2	605	100.0	801 7 AD35076
3	598	98.8	303 4 AAB95802 Human bre
4	542	89.6	140 4 AAU16619 Human pro
5	542	89.6	140 6 ABUS5688 Human nov
6	542	89.6	235 6 AAU16206 Human nov
7	542	89.6	235 6 ABUS5275 Human nov
8	500	82.6	688 3 AAV57898 Human tra
9	500	82.6	1390 5 ABUS65152 Human NOV
10	498	82.3	1937 4 ABUS8985 Drosophil
11	459	75.9	947 2 AAU81168 Transcrip
12	459	75.9	947 2 AAU07114 WO9304365
13	379	62.6	573 6 ABR41326 Human DIT
14	281.5	46.5	513 4 ABR68341 Drosophil
15	269	44.5	330 6 ABR41589 Human DIT
16	237	39.2	648 4 ABR63109 Drosophil
17	221.5	36.6	703 5 ABR93466 Herbicida
18	208	34.4	743 3 AAG52556 Arabidops
19	208	34.4	764 3 AAG52555 Arabidops
20	208	34.4	813 3 AAG52554 Arabidops
21	207	34.2	124 3 AAB32659 Eucalyptu
22	204	33.7	270 3 AAG42245 Arabidops
23	204	33.7	276 3 AAG42244 Arabidops
24	204	33.7	386 3 AAG42243 Arabidops
25	204	33.7	386 6 ABP70522 amino aci

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

CC interaction between the agent and the NAM or the expression product as a
 CC determination of the disorder. The products and methods can be used in
 CC the diagnosis, monitoring, research, or treatment of conditions
 CC characterised by the expression of various cancer associated antigens.
 CC The invention provides nucleic acid sequences and encoded polypeptides
 CC which are cancer associated antigen precursors expressed in human breast
 CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
 CC lung cancer
 XX
 SQ Sequence 754 AA;

Query Match 100.0%; Score 605; DB 2; Length 754;
 Best Local Similarity 100.0%; Pred. No. 2.8e-68;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 60
 Db 24 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 83
 QY 61 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 Db 84 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 134

RESULT 2
 ADC35076
 ID ADC35076 standard; protein; 801 AA.
 XX
 AC ADC35076;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human breast cancer antigen seq id 42.
 XX
 KW breast cancer; breast cancer diagnosis; breast cancer antigen.
 XX
 OS Homo sapiens.
 XX
 FN US2003108888-A1.
 XX
 PD 12-JUN-2003.
 XX
 PF 15-MAY-2002; 2002US-00146473.
 XX
 PR 15-MAY-2001; 2001US-0291150P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Scanlan MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;
 XX
 DR WPI; 2003-829397/77.
 DR N-PSDB; ADC35118.
 XX
 PT Diagnosing breast cancer in subject by obtaining biological sample from
 PT subject, contacting sample with breast cancer-associated polypeptides,
 PT determining specific binding between polypeptides and agents in sample.
 XX
 PS Example 2; SEQ ID NO 42; 173pp; English.
 XX
 CC The invention describes a method of diagnosing breast cancer in subject
 CC comprising contacting biological sample from subject with at least two
 CC different breast cancer-associated polypeptides (I) encoded by nucleic
 CC acid molecules (II) comprising sequence chosen from 42 fully defined
 CC sequences as given in specification, determining specific binding between
 CC (I) and agents in sample, where presence of the binding is diagnostic for
 CC breast cancer. The method is useful for diagnosing breast cancer in a
 CC subject. The sample is blood, lymph node fluid or breast discharge fluid.
 CC This is the amino acid sequence of a breast cancer antigen.
 XX
 SQ Sequence 801 AA;
 Query Match 100.0%; Score 605; DB 7; Length 801;
 Best Local Similarity 100.0%; Pred. No. 3.1e-68;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 60
 Db 24 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 83
 QY 61 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 Db 84 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 134

Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 60
 Db 71 KPGRTVNLQYLHKVVMKALWKHQFAPFPQVDAVKLGLPDYHKKIKQPMDMGTIKRRL 130
 QY 61 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
 Db 131 ENNYWAAASECMQDNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 181
 RESULT 3
 AAB95802
 ID AAB95802 standard; protein; 303 AA.
 XX
 AC AAB95802;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:18783.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 18783; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 303 AA;
Query Match 98.8%; Score 598; DB 4; Length 303;
Best Local Similarity 99.1%; Pred. No. 6.7e-68;
Matches 110; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KPGRTNQLYVHKVVMKALWKHQFAPFQPDVAVKLGLEDYHKIKIQPMDMTIKRRL 60
DB 74 KPGRTNQLYVHKVVMKALWKHQFAPFQPDVAVKLGLEDYHKIKIQPMDMTIKRRL 133
QY 61 ENNYWAAASECMQDFTNTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 134 ENNYWAAASECMQDFTNTFTNCYIYNKPTDDIVLMAQTPEKIFLOKVASMP 184
RESULT 4
AAU16619
ID AAU16619 standard; protein; 140 AA.
XX AC AAU16619;
DT 07-NOV-2001 (first entry)
XX DE Human novel secreted protein, Seq ID 1572.
XX KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX PN WO200155322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US0001341.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216847P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241231P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.

08-NOV-2000; 2000US-0246611P.
 08-NOV-2000; 2000US-0246613P.
 17-NOV-2000; 2000US-0249207P.
 17-NOV-2000; 2000US-0249208P.
 17-NOV-2000; 2000US-0249209P.
 17-NOV-2000; 2000US-0249210P.
 17-NOV-2000; 2000US-0249211P.
 17-NOV-2000; 2000US-0249212P.
 17-NOV-2000; 2000US-0249213P.
 17-NOV-2000; 2000US-0249214P.
 17-NOV-2000; 2000US-0249215P.
 17-NOV-2000; 2000US-0249216P.
 17-NOV-2000; 2000US-0249217P.
 17-NOV-2000; 2000US-0249218P.
 17-NOV-2000; 2000US-0249219P.
 17-NOV-2000; 2000US-0249220P.
 17-NOV-2000; 2000US-0249221P.
 17-NOV-2000; 2000US-0249222P.
 17-NOV-2000; 2000US-0249223P.
 17-NOV-2000; 2000US-0249224P.
 17-NOV-2000; 2000US-0249225P.
 17-NOV-2000; 2000US-0249226P.
 17-NOV-2000; 2000US-0249227P.
 17-NOV-2000; 2000US-0249228P.
 17-NOV-2000; 2000US-0249229P.
 17-NOV-2000; 2000US-0249230P.
 01-DEC-2000; 2000US-0250160P.
 01-DEC-2000; 2000US-0250391P.
 05-DEC-2000; 2000US-0251030P.
 05-DEC-2000; 2000US-0251988P.
 06-DEC-2000; 2000US-0256713P.
 08-DEC-2000; 2000US-0251479P.
 08-DEC-2000; 2000US-0251856P.
 08-DEC-2000; 2000US-0251868P.
 08-DEC-2000; 2000US-0251869P.
 08-DEC-2000; 2000US-0251989P.
 08-DEC-2000; 2000US-0251990P.
 11-DEC-2000; 2000US-0254097P.
 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-488783/53.
 N-PSDB; AAS26606.

New nucleic acid molecules encoding 461 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions and
 used as food additives or preservatives.

Claim 11; SEQ ID NO 1572; 980pp; English.

The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 in diagnosing a pathological condition or susceptibility to a
 pathological condition. Antibodies to the proteins can also be used in
 alleviating symptoms associated with the disorders and in diagnostic
 immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
 (ELISA). Disorders which are diagnosed or treated include autoimmune
 diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 nervous system disorders e.g. Alzheimer's disease, infections caused by
 bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 and many other disorders listed in the specification. The polypeptides
 can also be used to aid wound healing and epithelial cell proliferation,
 to prevent skin aging due to sunburn, to maintain organs before
 transplantation, for supporting cell culture of primary tissues, to
 regenerate tissues and in chemotaxis. The polypeptides can also be used
 as a food additive or preservative to increase or decrease storage
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 minerals, cofactors and other nutritional components. The present
 sequence represents a novel secreted protein of the invention. Note: The
 sequence data for this patent did not form part of the printed

Query Match 89.6%; Score 542; DB 4; Length 140;
 Best Local Similarity 87.4%; Pred. No. 3.7e-61; Indels 0; Gaps 0;
 Matches 97; Conservative 5; Mismatches 9;
 QY 1 KPGRTVNLQYLHKVYMKALWKHQFAPFPQPDVAVKGLPDYHKIKQPMDMGTIKRL 60
 Db 5 KPGRTVNLQYMNQVNVVTKLWKHQFAPFPQPDVAVKGLPDYHKIKQPMDMGTIKRL 64
 QY 61 ENNYWASACQDFTMTNCTYIYNKPTDDIVLNAQLEKIFLOKVASMP 111
 Db 65 ENNYWASACQDFTMTNCTYIYNKPTDDIVLNAQLEKIFLOKVASMP 115
 RESULT 5
 ABUS5688
 ID ABUS5688 standard; protein; 140 AA.
 XX
 AC ABUS5688;
 XX
 DT 18-MAR-2003 (first entry)
 XX
 DE Human novel polypeptide #775.
 XX
 KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.
 XX
 XX US2002132753-A1.
 XX 19-SEP-2002.
 XX
 XX 17-JAN-2001; 2001US-00764864.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 XX 04-FEB-2000; 2000US-0180628P.
 XX 28-JUN-2000; 2000US-0214886P.
 XX 07-JUL-2000; 2000US-0216647P.
 XX 07-JUL-2000; 2000US-0216880P.
 XX 11-JUL-2000; 2000US-0217487P.
 XX 11-JUL-2000; 2000US-0217496P.
 XX 14-JUL-2000; 2000US-0218290P.
 XX 26-JUL-2000; 2000US-0220963P.
 XX 14-AUG-2000; 2000US-0224518P.
 XX 14-AUG-2000; 2000US-0224519P.
 XX 14-AUG-2000; 2000US-0225267P.
 XX 14-AUG-2000; 2000US-0225268P.
 XX 14-AUG-2000; 2000US-0225270P.
 XX 14-AUG-2000; 2000US-0225447P.
 XX 14-AUG-2000; 2000US-0225757P.
 XX 22-AUG-2000; 2000US-0225758P.
 XX 14-AUG-2000; 2000US-0226868P.
 XX 30-AUG-2000; 2000US-0228924P.
 XX 01-SEP-2000; 2000US-0229287P.
 XX 01-SEP-2000; 2000US-0229343P.
 XX 01-SEP-2000; 2000US-0229344P.
 XX 01-SEP-2000; 2000US-0229345P.
 XX 05-SEP-2000; 2000US-0229509P.
 XX 05-SEP-2000; 2000US-0229513P.
 XX 08-SEP-2000; 2000US-0231413P.
 XX 21-SEP-2000; 2000US-0234223P.
 XX 21-SEP-2000; 2000US-0234274P.
 XX 25-SEP-2000; 2000US-0234997P.
 XX 27-SEP-2000; 2000US-0235834P.
 XX 29-SEP-2000; 2000US-0236327P.
 XX 29-SEP-2000; 2000US-0236367P.
 XX 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX
DR WPI; 2003-147444/14.
DR N-PSDB; ABX73947.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 11; SEQ ID NO 1572; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
XX Sequence 140 AA;
XX
Query Match 89.6%; Score 542; DB 6; Length 140;
Best Local Similarity 87.4%; Pred. No. 3.7e-61;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
Oy 1 KPCRVTNQYLHKYMKALWKEQFAWPPQVDVAVKGLPDYHKTIKOPMDMGTIKKRL 60
Db 5 KPCRKTNQLQYQNVVVKLWKEQFAWPPQVDVAVKGLPDYHKTIKOPMDMGTIKKRL 64
Oy 61 ENNYWAASECMDQFNTMTFNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 65 ENNYWASSECMDQFNTMTFNCYIYNKPTDDIVLMAQALEKIFLQKVQAMP 115
RESULT 6
AAU16206
ID AAU16206 standard; protein; 235 AA.
XX
AC AAU16206;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1159.

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
PF 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2000US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-488783/53.
XX N-PSDB; AAS26193.
DR
XX
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1159; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunosays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed
Query Match 89.6%; Score 542; DB 4; Length 235;
Best Local Similarity 87.4%; Pred. No. 7, 4e-61;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLKVKVMKALWKHGFAMFPFQPDVAVKLGFLPDYHKIKQPMDMGTIKRL 60
Db 5 KPGRTNQLQYMNQVVKTLWKHGFAMFPFQPDVAVKLGFLPDYHKIKQPMDMGTIKRL 64
QY 61 ENNYWASACQDFNTMTNCTYINPKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 65 ENNYWASACQDFNTMTNCTYINPKPTDDIVLMAQTLEKIFLQKVASMP 115
RESULT 7
ABUS5275
ID ABUS5275 standard; protein; 235 AA.
XX
AC ABUS5275;
XX
XX 18-MAR-2003 (first entry)
DT
XX Human novel polypeptide #362.
DE
XX Human; neural disorder; immune system disorder; renal disorder;
KW Human; muscular disorder; respiratory disease; reproductive disorder;

KW Gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX
 OS Homo sapiens.

XX US2002132753-A1.

XX 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 28-JUN-2000; 2000US-0214886P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216680P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

XX 14-AUG-2000; 2000US-0225758P.

XX 22-AUG-2000; 2000US-0226686P.

XX 30-AUG-2000; 2000US-0228924P.

XX 01-SEP-2000; 2000US-0229287P.

XX 01-SEP-2000; 2000US-0229343P.

XX 01-SEP-2000; 2000US-0229344P.

XX 01-SEP-2000; 2000US-0229345P.

XX 05-SEP-2000; 2000US-0229509P.

XX 05-SEP-2000; 2000US-0229513P.

XX 08-SEP-2000; 2000US-0231413P.

XX 21-SEP-2000; 2000US-0234223P.

XX 21-SEP-2000; 2000US-0234274P.

XX 25-SEP-2000; 2000US-0234997P.

XX 27-SEP-2000; 2000US-0235834P.

XX 29-SEP-2000; 2000US-0236327P.

XX 29-SEP-2000; 2000US-0236367P.

XX 29-SEP-2000; 2000US-0236368P.

XX 29-SEP-2000; 2000US-0236369P.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX
 PS Claim 11; SEQ ID NO 1159; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention

XX SQ Sequence 235 AA;

Query Match 89.6%; Score 542; DB 6; Length 235;

Best Local Similarity 87.4%; Pred. No. 7.4e-61;

Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVVKALWKQFAWFPQPDVAVKGLPDYHKIIPQMDMGTIKRL 60

DB 5 KPRKTNQLQYMNQVVKTLWKQFAWFPYQPDVAKIKNLPDYHKIIPQMDMGTIKRL 64

QY 61 ENNYWASCECMQDFTMTNCTYLYNKPTDDIVLMAQTLEKIPLQKVASMP 111

DB 65 ENNYWASCECMQDFTMTNCTYLYNKPTDDIVLMAQTLEKIPLQKVASMP 115

RESULT 8

AAV57898

ID AAV57898 standard; protein; 688 AA.

XX AAV57898;

XX 23-MAR-2000 (first entry)

XX Human transmembrane protein HTPN-22.

XX Human; transmembrane protein; HTPN; diagnosis; immunospecific;
 KW antiproliferative; neuroprotective; immune disorder;
 KW reproductive disorder; smooth muscle disorder; neurological disorder;
 KW gastrointestinal disorder; developmental disorder;
 KW cell proliferative disorder.

XX Homo sapiens.

XX WO9961471-A2.

XX 02-DEC-1999.

XX 28-MAY-1999; 99WO-US011904.

XX 29-MAY-1998; 98US-0087260P.

XX 02-JUL-1998; 98US-0091674P.

XX 02-OCT-1998; 98US-0102954P.

XX 24-NOV-1998; 98US-0109869P.

XX (INCY-) INCYTE PHARM INC.

XX Tang YT, Lal P, Hillman JL, Yue H, Guegler KJ, Corley NC;

PI Bandman O, Patterson C, Gorgone GA, Kaser MR, Baughn MR, Au-Young J;

XX

XX (ROSE/) ROSEN C A.
 XX (RUBE/) RUBEN S M.
 XX (BARA/) BARASH S C.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI; 2003-147444/14.
 XX N-PSDB; ABX73534.

DR WPI; 2000-072605/06.
 DR N-PSDB; AA256719.
 XX
 PT Proteins, polynucleotides, vectors, host cells and antibodies used to
 PT diagnose, treat or prevent immune, reproductive, smooth muscle,
 PT neurological, gastrointestinal, developmental and cell proliferative
 PT disorders.
 XX
 PS Claim 1; Page 128-130; 229pp; English.
 XX
 CC AA256698 to AA256776 encode AAY57877 to AAY57955 which represent human
 CC transmembrane proteins designated HTMPN-1 to HTMPN-79, respectively. The
 CC transmembrane protein have immunospecific, antiproliferative and
 CC neuroprotective activities. The human transmembrane proteins,
 CC polynucleotides encoding them and other compositions and methods from the
 CC present invention, can be used for the diagnosis, treatment or prevention
 CC of immune, reproductive, smooth muscle, neurological, gastrointestinal,
 CC developmental and cell proliferative disorders. The HTMPN's can be used
 CC to treat or prevent disorders associated with a decreased expression or
 CC activity of HTMPN
 CC
 SQ Sequence 688 AA;
 Query Match 82.6%; Score 500; DB 3; Length 688;
 Best Local Similarity 80.2%; Pred. No. 7.6e-55;
 Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
 QY 1 KPGRTNQLVHLKVMKALWKHOFAPFPQPDVAVKGLPDYHKIKQMDMGTTIKRRL 60
 DB 55 KPGRTNQLVHLVRLVLTWKHOFAPFPQPDVAVKGLPDYHKIKQMDMGTTIKRRL 114
 QY 61 ENNYWAAECQDPTMTNCTYINXPTDDIVLMAOTLEKIFLOKVASMP 111
 DB 115 ENNYWAAECIQDPTMTNCTYINXPTDDIVLMAOTLEKIFLOKVASMP 165
 RESULT 9
 ABU65152
 ID ABU65152 standard; protein; 1390 AA.
 XX
 AC ABU65152;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV79a protein.
 XX
 KW NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US006908.
 XX
 PR 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.

PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0277338P.
 PR 30-MAR-2001; 2001US-0279959P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 31-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 14-NOV-2001; 2001US-0333272P.
 PR 21-NOV-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 03-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 07-MAR-2002; 2002US-00092900.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Patturajan M, Gangolli E, Vernet CM, Guo X, Tchernev V;
 PI Fernandes ER, Casman SU, Malyankar UM, Gerlach V, Liu Y, Anderson D;
 PI Spaderina SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
 PI Lepley DM, Rieger DK;
 XX
 XX WPI; 2002-723332/78.
 DR N-PSDB; ABX97119.
 XX
 PT NOVX polypeptides and polynucleotides, useful for preventing or treating
 PT a disorder associated with aberrant NOVX expression or activity e.g.,
 PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
 PT asthma.
 XX
 PS Claim 1; Page 323; 1103pp; English.
 XX
 CC This invention describes novel human NOVX polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
 CC activity. Pharmaceutical compositions comprising the NOVX proteins or
 CC nucleic acid molecules or NOVX antibodies are useful for preventing or
 CC treating a disorder associated with aberrant NOVX expression or activity

AA070114
ID AA070114 standard; protein; 947 AA.

AC AA070114;

XX 02-JUL-1999 (first entry)

DE WO9904265 Seq ID No: 685.

KW Cancer associated antigen; diagnosis; research; treatment; human;

KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;

XX prostate cancer.

OS Homo sapiens.

PN WO9904265-A2.

PD 28-JAN-1999.

PF 15-JUL-1998; 98WO-US014679.

PR 17-JUL-1997; 97US-00896164.

PR 10-OCT-1997; 97US-0061599P.

PR 10-OCT-1997; 97US-0061765P.

PR 11-OCT-1997; 97US-00948705.

PR 22-JUN-1998; 97GB-00021657.

PR 22-JUN-1998; 98US-00102322.

XX (LUDW-) LUDWIG INST CANCER RES.

PI Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;

PI O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;

XX WPI; 1999-132448/11.

XX New isolated cancer associated nucleic acids and polypeptides - isolated

PT using sera from cancer patients, used to develop products for the

PT diagnosis, monitoring or treatment of cancers.

XX Disclosure; Page 728-730; 787pp; English.

XX The invention relates to a method for diagnosing a disorder characterised

CC by expression of a human cancer associated antigen precursor coded for by

CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a

CC biological sample isolated from a subject with an agent that specifically

CC binds to the NAM, an expression product or a fragment of an expression

CC product complexed with an HLA molecule; and (b) determining the

CC interaction between the agent and the NAM or the expression product as a

CC determination of the disorder. The products and methods can be used in

CC the diagnosis, monitoring, research, or treatment of conditions

CC characterised by the expression of various cancer associated antigens.

CC The invention provides nucleic acid sequences and encoded polypeptides

CC which are cancer associated antigen precursors expressed in human breast

CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and

XX lung cancer

ABR41326 standard; protein; 573 AA.

ABR41326;

02-JUN-2003 (first entry)

Human DITHP intracellular signalling protein.

Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;

KW cancer; cell proliferative disorder; autoimmune disorder;

KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;

KW neurological disorder; gastrointestinal disorder; transport disorder;

KW connective tissue disorder; drug screening; proteome analysis;

KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;

KW disease model; toxicological testing; transcript imaging;

KW intracellular signalling.

OS Homo sapiens.

WO200297031-A2.

05-DEC-2002.

27-MAR-2002; 2002WO-US010056.

28-MAR-2001; 2001US-0279619P.

29-MAR-2001; 2001US-0280067P.

29-MAR-2001; 2001US-0280068P.

16-MAY-2001; 2001US-0291280P.

17-MAY-2001; 2001US-0291829P.

19-JUN-2001; 2001US-0291849P.

20-JUN-2001; 2001US-0299428P.

20-JUN-2001; 2001US-0299776P.

20-JUN-2001; 2001US-0300001P.

(INCY-) INCYTE GENOMICS INC.

Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;

Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;

Feratta CH, David M, Lewis SA, Chen AJ, Panzer SR, Harris B;

Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

WPI; 2003-129518/12.

N-PSDB; ACC46266.

Novel human diagnostic and therapeutic polypeptide useful for identifying

PT test compound which specifically binds to a polypeptide encoded by human

PT diagnostic and therapeutic polynucleotide, and to induce antibodies.

XX Claim 27; SEQ ID NO 861; 591pp; English.

The invention relates to novel human diagnostic and therapeutic

CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded

CC proteins (DITHP; ABR41326-ABR41812). The invention also relates to

Query Match 75.9%; Score 459; DB 2; Length 947;

Best Local Similarity 73.0%; Pred. No. 2.1e-49;

Matches 81; Conservativity 16; Mismatches 14; Indels 0; Gaps 0;

QY 1 KPGRTNQLQVHLKVMKALWKFQFPPFQPDVAVKLGIPDYHKIKQPMWGKTKRRL 60

DB 24 KNGRLTNQLQVHLKVMKALWKFQFPPFQPDVAVKLGIPDYHKIKQPMWGKTKRRL 83

QY 61 ENNYWAASECQFNTMTNCTYINXPTDDIVLMAQTLKIFLQKVASNP 111

DB 84 ENKYAKASACIEDNTWFSNCLYNKFGDDIVLMAQALEKLFQKLSQNP 134

RESULT 13

ABR41326

CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a DITHP protein which has intracellular
CC signalling activity. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 573 AA;

Query Match 62.6%; Score 379; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 2e-39;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 DYHKIIKQPMXGTTIKRLENNYWAASECMQDFTMTNCYIYNKPTDDIVLMAQTLEK 101
DB 5 DYHKIIKQPMXGTTIKRLENNYWAASECMQDFTMTNCYIYNKPTDDIVLMAQTLEK 64
QY 102 IFLOKVASMP 111
DB 65 IFLOKVASMP 74

RESULT 14
ABB68341
ID ABB68341 standard; protein: 513 AA.
XX
AC ABB68341;
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 31815.

XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX Drosophila melanogaster.
XX WO200171042-A2.

XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL12444.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX Disclosure; SEQ ID NO 31815; 2lpp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLU6176-ABLU30511), expressed DNA
XX sequences (ABLU1840-ABLU16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 513 AA;

Query Match 46.5%; Score 281 5; DB 4; Length 513;
Best Local Similarity 44.5%; Pred. No. 5.8e-27;
Matches 49; Conservative 25; Mismatches 35; Indels 1; Gaps 1;
QY 1 KPGRTVITQQLVHLKVMKALMKHQFAWFFRQPDVAVKGLPVDYHKIIKQPMXGTTIKRRL 60
DB 36 RPRGRTWILBEL-KSVLNCLEWRNRFVYHFRPVDVSLGVDPYHAVVKHPMDLSTIKRRL 94
QY 61 ENNYWAASECMQDFTMTNCYIYNKPTDDIVLMAQTLEKIFLOKVASM 110
DB 95 HNKYYWQASEALEDFKLIFDNCLLYNLEGSFVYQAGKLLMEAFYMRMESI 144

RESULT 15
ABR41589
ID ABR41589 standard; protein: 330 AA.
XX
AC ABR41589;
DT 02-JUN-2003 (first entry)
XX
DE Human DITHP antigen recognition protein.

XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
XX cancer; cell proliferative disorder; autoimmune disorder;
XX inflammatory disorder; infection; hormonal disorder; metabolic disorder;
XX neurological disorder; gastrointestinal disorder; transport disorder;
XX connective tissue disorder; drug screening; proteome analysis;
XX gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
XX disease model; toxicological testing; transcript imaging;
XX antigen recognition.

XX Homo sapiens.
XX WO200297031-A2.
XX
PD 05-DEC-2002.

XX 27-MAR-2002; 2002WO-US010056.
XX
XX 28-MAR-2001; 2001US-0279619P.
XX 29-MAR-2001; 2001US-0280067P.
XX 29-MAR-2001; 2001US-0280068P.
XX 16-MAY-2001; 2001US-0291280P.
XX 17-MAY-2001; 2001US-0291849P.
XX 17-MAY-2001; 2001US-0291849P.
XX 19-JUN-2001; 2001US-0299428P.
XX 20-JUN-2001; 2001US-0299776P.
XX 20-JUN-2001; 2001US-0300001P.
XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX WPI; 2003-129518/12.
XX N-PSDB; ACC46527.

XX Novel human diagnostic and therapeutic polypeptide useful for identifying
XX test compound which specifically binds to a polypeptide encoded by human
XX diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX Claim 27; SEQ ID NO 1124; 59lpp; English.

XX The invention relates to novel human diagnostic and therapeutic
XX polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
XX proteins (DITHP; ABR41136-ABR41812). The invention also relates to
XX polynucleotide sequences at least 90% identical to the dithp cDNA
XX sequences of the invention; recombinant vectors, host cells and

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:04:05 ; Search time 21 Seconds
(without alignments) 508.441 Million cell

Title: US-09-784-553C-19
 Perfect score: 605
 Sequence: 1 KPGAVTNQLQVLHKVVMKAL.....IVLMAQTLEKIFLOKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

```
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SIMMARTES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	605	100.0	754	2	A56619	female sterile hom	
2	594	98.2	733	2	T28145	RING3 kinase - chi	
3	498	82.3	2038	2	A43742	female sterile hom	
4	295	48.9	1250	2	T22845	hypothetical prote	
5	279	46.1	1087	2	T22847	hypothetical prote	
6	221.5	36.6	703	2	T48600	kinase-like protei	
7	215	35.5	578	2	T40984	transcription fact	
8	204	33.7	400	2	T00472	probable RING3 pro	
9	201	33.2	461	2	D96757	hypothetical prote	
10	198	32.7	832	1	S71788	P/CAF protein - hu	
11	195.5	32.3	374	2	T33328	hypothetical prote	
12	195	32.2	678	2	T49584	bromodomain protei	
13	190.5	31.5	361	2	T42517	bromodomain protei	
14	188	31.1	766	2	A86198	hypothetical prote	
15	187	30.9	440	2	H86312	F2H15.2 protein.	
16	184.5	30.5	2441	2	S39161	CREB-binding prote	
17	183.5	30.3	638	2	S67605	hypothetical prote	
18	182	30.1	439	1	T28051	transcription fact	
19	181.5	30.0	2414	2	A54377	transcription adap	
20	180.5	29.8	2440	2	S39162	transcription coac	
21	179.5	29.7	586	2	T47620	histon acetyltrans	
22	176	29.1	405	2	T21433	hypothetical prote	
23	176	29.1	432	2	T21435	hypothetical prote	
24	176	29.1	510	2	T21430	hypothetical prote	
25	176	29.1	686	2	S55955	bromodomain protei	
26	172	28.4	454	2	T37933	transcription acti	
27	168.5	27.9	763	2	E96613	hypothetical prote	
28	166	27.4	3130	2	T13828	CREB-binding prote	
29	162	26.8	2027	2	S60123	hypothetical prote	

ALIGNMENTS

RESULT 1

A56619 female sterile homeotic (fsh) homolog RING3 - human
C:Species: Homo sapiens (man)
C:date: 21-Jul-1995 #sequence_revision 03-Aug-1995
C:Accession: A56619; S18860; S40781
R:Beck, S.; Hanson, I.; Kelly, A.; Pappin, D.J.; Tre
DNA Seq. 2, 203-210, 1992
A:title: A homologue of the *Drosophila* female steri
A:Reference number: A56619; WUID:92329974; PMID:135

A;Accession: A58619
A;Accession: A58619
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-754 <BEC>
A;Cross-references: EMBL:X62083; NID:g31471; PID:AAA68990.1; PID:g31472; EMBL:M80613; N
A;Note: sequence extracted from NCBI backbone (NCBIP:108781)
A;Note: sequence extracted from NCBI backbone (NCBIP:108781)

A;Gene: RING3
C-Superfamily: unassigned
bromodomain proteins: bromodomain homology

C: Keywords: duplication
C/superfamiliarity: unassigned promodomain problem

C/Keywords: duplication
F: 52-109/Domain: bromodomain homology <BRO1>

F:325-382/Domain: bromodomain homology <BRO2>

Query Match 100.0%; Score 605; DB 2; Length 754;

Method	Best Local Similarity	Pred. No.	7.6e-58
Matches 111	Conservative	0	Mismatches 0
Indels	0	0	Gaps 0

1 KPGRVTNOLOYLHKVVMKALWKHQFAWPFROQVDAVKLGLPDYHKIKQPMDMGTIKRRL 60

24 KPRVYNOLYUHKVVMKALWKHOFAPWFRQVDAVKLGLPDYHKILKOPMDMGITKEEL 83

61 ENNVYWAASECMODENTMTFTNCYIYNKPTDDIVLMAOTLEKIFLOKVAMP 111

QY	SNNIWWAASECQDINMILTIRCLIANKEIDDIIVETFEQLSEKTLTQVASHII
Dh	NNVVWAAASECMODENTMTFTNCVTYNKPBTDDIVLMACLEKIILOKVASMP

RESULT 2

T28145

RING3 kinase - chicken

C;Species: Gallus gallus (chicken)

C;Date: 15-Oct-1999

C;Accession: T28145
P.Milne & Kaufman J. Beck S

R; Milne, S.; Kaurman, J.; Beck, S. submitted to the EMBL Data Library. May 1998

A;Description: DNA sequencing and analysis of the chicken major histocompatibility complex submitted to the EMBL data library, May 1990
 1.Reference number: 720475

A;Reference number: Z20475
A:Accession: T28145

A;Accession: 12013
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-733 <MIL>

A;Cross-references: EMBL:AL023516; NID:e1292539; PID:e1292549; PIDN:CAA18965.1

A; Experimental source: clone CB12

submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19625
 A:Accession: T22845
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1250 <WIL>
 A:Cross-references: EMBL:Z69646; PIDN:CAA93473.1; GSPDB:GN00028; CESP:F57C7.1a
 A:Experimental source: clone F57C7
 C:Genetics:
 A:Gene: CESP:F57C7.1a
 A:Map position: X
 A:Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3
 C:Superfamily: bromodomain homology
 F:307-364/Domain: bromodomain homology <BRO1>
 F:619-676/Domain: bromodomain homology <BRO2>

Query Match 48.9%; Score 296; DB 2; Length 1250;
 Best Local Similarity 46.8%; Pred. No. 7e-24;
 Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFANPFRQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
 DB 279 KPTRHTNCLDFVLFVTVKDKLKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAASECQDNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
 DB 339 RNLYWCAEDAIKDINQVFINCYSFNPPEYDVYKMAKTLKQVLSQLTQLP 399

RESULT 5
 T22847
 hypothetical protein F57C7.1b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T22847
 R:White, S.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19625
 A:Accession: T22847
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1087 <WIL>
 A:Cross-references: EMBL:Z69646; PIDN:CAA93475.1; GSPDB:GN00028; CESP:F57C7.1b
 A:Experimental source: clone F57C7
 C:Genetics:
 A:Gene: CESP:F57C7.1b
 A:Map position: X
 A:Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1
 C:Superfamily: bromodomain homology
 F:307-364/Domain: bromodomain homology <BRO1>
 F:579-636/Domain: bromodomain homology <BRO2>

Query Match 46.1%; Score 279; DB 2; Length 1087;
 Best Local Similarity 44.1%; Pred. No. 4.3e-22;
 Matches 49; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFANPFRQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
 DB 279 KPTRHTNCLDFVLFVTVKDKLKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAASECQDNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
 DB 339 RNLYWCAEDAIKDINTLFDNCKFNDRNDIYMCENIEGVVGQGLEWMP 399

RESULT 6
 T48600
 kinase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F18022.60
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
 C:Accession: T48600
 R:Sevan, M.; Murphy, G.; Ridley, G.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lea

submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19625
 A:Accession: T22845
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1250 <WIL>
 A:Cross-references: EMBL:Z69646; PIDN:CAA93473.1; GSPDB:GN00028; CESP:F57C7.1a
 A:Experimental source: clone F57C7
 C:Genetics:
 A:Gene: CESP:F57C7.1a
 A:Map position: X
 A:Introns: 262/3; 351/3; 391/3; 431/3; 667/3; 835/2; 997/1; 1127/1; 1218/3
 C:Superfamily: bromodomain homology
 F:307-364/Domain: bromodomain homology <BRO1>
 F:619-676/Domain: bromodomain homology <BRO2>

Query Match 48.9%; Score 296; DB 2; Length 1250;
 Best Local Similarity 46.8%; Pred. No. 7e-24;
 Matches 52; Conservative 23; Mismatches 36; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFANPFRQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
 DB 279 KPTRHTNCLDFVLFVTVKDKLKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAASECQDNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
 DB 339 RNLYWCAEDAIKDINQVFINCYSFNPPEYDVYKMAKTLKQVLSQLTQLP 399

RESULT 5
 T22847
 hypothetical protein F57C7.1b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
 C:Accession: T22847
 R:White, S.
 submitted to the EMBL Data Library, February 1996
 A:Reference number: Z19625
 A:Accession: T22847
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-1087 <WIL>
 A:Cross-references: EMBL:Z69646; PIDN:CAA93475.1; GSPDB:GN00028; CESP:F57C7.1b
 A:Experimental source: clone F57C7
 C:Genetics:
 A:Gene: CESP:F57C7.1b
 A:Map position: X
 A:Introns: 262/3; 351/3; 391/3; 627/3; 795/2; 957/1
 C:Superfamily: bromodomain homology
 F:307-364/Domain: bromodomain homology <BRO1>
 F:579-636/Domain: bromodomain homology <BRO2>

Query Match 46.1%; Score 279; DB 2; Length 1087;
 Best Local Similarity 44.1%; Pred. No. 4.3e-22;
 Matches 49; Conservative 24; Mismatches 38; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFANPFRQPDVAVKLGDPDYHKIILKQPMDMGTIKRRL 60
 DB 279 KPTRHTNCLDFVLFVTVKDKLKHSHWPFQLPVDAIKLEIPEYHNI VNTPMDLSTIKRRL 338

QY 61 ENNYWAASECQDNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
 DB 339 RNLYWCAEDAIKDINTLFDNCKFNDRNDIYMCENIEGVVGQGLEWMP 399

RESULT 6
 T48600
 kinase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F18022.60
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 02-Sep-2000
 C:Accession: T48600
 R:Sevan, M.; Murphy, G.; Ridley, G.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lea

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493
A:Accession: T48600
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-703 <BEV>
A:Cross-references: EMBL:DB
A:Experimental source: cultivar Columbia; BAC clone F18022
C:Genetics:
A:Map position: 5
A:Introns: 370/3; 387/1; 423/1; 461/3; 484/1; 511/1; 549/3; 572/1; 619/3
A:Note: F18022.60
C:Superfamily: bromodomain homology
F:157-214/Domain: bromodomain homology <BRO>

Query Match 36.6%; Score 221.5; DB 2; Length 703;
Best Local Similarity 41.5%; Pred. No. 4.7e-16;
Matches 44; Conservative 15; Mismatches 44; Indels 3; Gaps 1;

QY 1 KGRVTNQLQY-LHK---VVKALWKHQFAMPPRQPDVAVKGLPDYHKIIKQPMDMGTHK 57
DB 126 KPTPTVAVMLKMQCEALKRLMHSQYGVFNTFVVKLNILDFNVNIEHPMDLGTVK 185

QY 58 RLENNYWAASECMQDFMTNCTVYVYKPTDDIVLMAQTLEKIF 103
DB 186 NKLTSCTYSCPSEFAADVRLTFSNANTYPPGNDVYVWADTLRKTF 231

RESULT 7

T40984
transcription factor bdf1 homolog SPC1450.02 - fission yeast (Schizosaccharomyces pombe)
N:Alternate names: protein SPC191.13
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Dec-2000
A:Accession: T40984; T41225
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21962
A:Accession: T40984
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-578 <LYN1>
A:Cross-references: EMBL:AL049559; NID:94581510; PIDN:CAB40169.1; PID:g4581512; GSPDB:GN
A:Experimental source: strain 972h; cosmid c1450
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21904
A:Accession: T41225
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-111 <LYN2>
A:Cross-references: EMBL:AL049644; NID:94678667; PIDN:CAB41059.1; PID:g4678679; GSPDB:GN
A:Experimental source: strain 972h; cosmid c191
C:Genetics: <C145>
A:Gene: SPDB:SPCC1450.02
A:Map position: 3
C:Genetics: <C191>
A:Gene: SPDB:SPCC191.13
A:Map position: 1
C:Superfamily: bromodomain homology
F:109-166/Domain: bromodomain homology <BRO1>
F:279-336/Domain: bromodomain homology <BRO2>

Query Match 35.5%; Score 215; DB 2; Length 578;
Best Local Similarity 40.9%; Pred. No. 1.9e-15;
Matches 47; Conservative 15; Mismatches 49; Indels 4; Gaps 2;

QY 1 KGRVTNQLQY-LHKVVKALWKHQ---FAMPPRQPDVAVKGLPDYHKIIKQPMDMGTH 56

DB 247 KPRRRKNSQMRFCSTVLKELYRQYSEFAFPYQPDVAVKGLPDYHKIIKQPMDMGTH 306

QY 57 KRLLENNYWAASECMQDFMTNCTVYVYKPTDDIVLMAQTLEKIFLOKVASMP 111

DB 307 QSKLNKNEYSTLEEFESDILLMFNNCTYNPPGTPVHVMGRQLENVFKKWEARP 361

RESULT 8

T00472
probable RING3 protein [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F1913.13
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T00472; C84762
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul,
submitted to the EMBL Data Library, April 1998
A:Description: Arabidopsis thaliana chromosome II BAC F1913 genomic sequence.
A:Reference number: Z14160
A:Accession: T00472
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-400 <ROU>
A:Cross-references: EMBL:AC004238; NID:g3033373; PID:g3033386
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; A.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
C.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84762
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <STO>
A:Cross-references: GB:AE002093; NID:g3033386; PIDN:AAAC12830.1; GSPDB:GN00139
C:Genetics:
A:Gene: At3g34900; F1913.13
A:Map position: 2
A:Introns: 6/3; 49/3; 120/3; 146/3; 218/3; 251/3; 271/2; 335/3
C:Superfamily: bromodomain homology
F:130-187/Domain: bromodomain homology <BRO>

Query Match 33.7%; Score 204; DB 2; Length 400;
Best Local Similarity 39.6%; Pred. No. 2e-14;
Matches 36; Conservative 26; Mismatches 29; Indels 0; Gaps 0;

QY 16 VKKALWKHQFAMPPRQPDVAVKGLPDYHKIIKQPMDMGTHKRLLENNYWAASECMQDF 75

DB 117 MFRQIAQHKWAPPELEFVDVKGGLHDYKVKYIEKPMDLGTHKKQESSEYSNVEIYADV 176

QY 76 NTMTNCTVYVYKPTDDIVLMAQTLEKIFLOK 106

DB 177 RLVPKRNAMRYNEKEEDVYVMAESLTERPEEK 207

RESULT 9

D36757
hypothetical protein T18x17.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D36757
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.D.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D36757
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-461 <STO>
A:Cross-references: GB:AE005173; NID:96598866; PIDN:AAF18720.1; GSPDB:GN00141

```
C;Superfamily: bromodomain homology
F;141-198/Domain: bromodomain homology <BRO>

Query Match      32.3%; Score 195.5; DB 2; Length 374;
Best Local Similarity 40.4%; Pred. No. 1.5e-13;
Matches 42; Conservative 12; Mismatches 42; Indels 9; Gaps 1;
```

QY 11 YLHKVYMAL-----WKHQFAMFPFROPVDVAVKLGLPDYHKKIQPMDMGTIKRRL 61
 ::|||::| |||||::| |||::| |||::| |||:
Db 114 LHLDLKLCLGILKEFKSTHDSTFTPFRRPVDWLGLTDYHEVIKKPMDMSTIRKKLI 173

QY 62 NNYWAAASECMODNTFTNCTIYNKPDTDDIVLMAQTLEKIFLOK 106
 ::|||::| |||||::| |||::| |||::| ||:
Db 174 GREYTAVEFRDKFLINNCLLYNNEGDPVADFALQFRKFAAK 218

RESULT 12
T49984
bromodomain protein-like - Arabidopsis thaliana
N;Alternate names: protein F12B17.100
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Sep-2000
C;Accession: T49984
R;Bevan, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z25026
A;Accession: T49984
A>Status: Preliminary
A:Molecule type: DNA
A;Residues: 1-678 <BEV>
A;Cross-references: EMBL:AL353995; GSPDB:GN00063; ATSP:F12B17.100
A;Experimental source: cultivar Columbia, BAC clone F12B17
C;Genetics:
A;Gene: ATSP:F12B17.100
A;Map position: 5
A;Intons: 73/3; 560/3; 629/1; 656/2
A;Introns: 73/3; 560/3; 629/1; 656/2
C;Superfamily: bromodomain homology
F;269-326/Domain: bromodomain homology <BRO>

Query Match 32.2%; Score 195; DB 2; Length 678;
Best Local Similarity 37.4%; Pred. No. 3.5e-13;
Matches 40; Conservative 17; Mismatches 38; Indels 12; Gaps 1;

QY 16 VNKAHWKHQFAMPFPQPVDVAVKLGLPDYHKKIQPMDMGTIKRRLNNYYNAECMQDF 75
 :: |||::| | | | | : : : : : : : : : : : : : : :
Db 256 ILVKLMKHKSWSPLNPVDVVGLGLDHYHRIDVPMDLTGVKNLEKGYLRSPIDFASDV 315

QY 76 NTMTFNCTIYNKPDTDDIVLMAQ-----TLEKIFLOKVASM 110
 ||| | | | : : : : : : : : : : : : : : : :
Db 316 RLTTNAMSYPNGQQDVLMAEKLQSFDVWFNPITLKRFEAQEYKVM 362

RESULT 13
T42517
bromodomain protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C;Accession: T42517
R;Yoshioaka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A>Title: Identification of open reading frames in Schizosaccharomyces pombe cdnas.
A;Reference number: Z17323; MUID:98162722; PMID:9501991
A;Accession: T42517
A>Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: mRNA
A;Residues: 1-361 <YOS>
A;Cross-references: EMBL:D89157; NID:g1749521; PIDN:BAA13819.1; PID:g1749522
A;Experimental source: strain PR745
C;Superfamily: bromodomain homology
F;64-121/Domain: bromodomain homology <BRO>

Query Match 31.5%; Score 190.5; DB 2; Length 361;
Best Local Similarity 40.4%; Pred. No. 5.2e-13;

Matches 40; Conservative 17; Mismatches 39; Indels 3; Gaps 1;
QY 16 VMKALWKHC--FANPFPQPVDAVKLGDPYHKIIKQPMDMGTIKRRLNNYVWAAASECM 72
Db 48 VLKELLKQHEAYAYPFYKVPNPTACGCPDYFKYKHPMDLGTMONKLNHNEYASMKAFE 107
QY 73 QDFNTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVASMP 111
bb 108 ADWLMLFKNCYKFNAGTGVHLMGKLESIFQKLWANKP 146

RESULT 14
A86198
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86198
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-766 <STO>
A:Cross-references: GB:AE005172; NID:g9844128; PIDN:AAF60220.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 31.1%; Score 188; DB 2; Length 766;
Best Local Similarity 43.0%; Pred. No. 2.4e-12;
Matches 40; Conservative 13; Mismatches 40; Indels 0; Gaps 0;
QY 16 VMKALWKHCFAWPFROPVDAVKLGDPYHKIIKQPMDMGTIKRRLNNYVWAAASECMQDF 75
Db 428 LLERLMKHKHGWFNAPVDVKGGLLDYTTIIEHPMDLGTIKSALMKNLYKSPREFAEADV 487
QY 76 NTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVA 108
Db 488 RLTFHNTYINPEGQDVHLMVATLLQIFERWA 520

RESULT 15
H86312
P2H15.2 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C:Accession: H86312
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <STO>
A:Cross-references: GB:AE005172; NID:g9665057; PIDN:AAF97259.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 30.9%; Score 187; DB 2; Length 440;
Best Local Similarity 38.9%; Pred. No. 1.6e-12;
Matches 37; Conservative 17; Mismatches 41; Indels 0; Gaps 0;
QY 16 VMKALWKHCFAWPFROPVDAVKLGDPYHKIIKQPMDMGTIKRRLNNYVWAAASECMQDF 75
Db 139 LLTKLMKHKSAWVFNVPVDKGLGHDYHNIVKEPMDLGTIVTKLKGSLYKSPDFAEDV 198
QY 76 NTMFTNCYIYNKFTDDIVLMAQTLEKIFLQKVASM 110
Db 199 RLTFNNAILNPIGHVDVYRFAELLNMFEDKXYSI 233

Search completed: July 9, 2004, 13:08:07
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:00:15 ; Search time 17 Seconds
(without alignments)
339.987 Million cell updates/sec

Title: US-09-784-553C-19

Perfect score: 605
Sequence: 1 KGRVTNQLYLHKVMKAL.....IVLMAQTLEKIFLOKVASMP 111

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	605	100.0	801	1 BRD2_HUMAN	P25440 homo sapien
2	542	89.6	726	1 BRD3_HUMAN	Q15059 homo sapien
3	500	82.6	1362	1 BRD4_HUMAN	O60885 homo sapien
4	498	82.3	2038	1 FSH_DROME	P13709 drosophila
5	202.5	33.5	727	1 YK82_SCHPO	Q9hgp4 schizosacch
6	187	30.9	832	1 PCAF_HUMAN	Q9h2831 homo sapien
7	184.5	30.5	2441	1 CBP_MOUSE	P45481 mus musculus
8	182	30.1	439	1 GCNE_YEAST	Q03330 saccharomyc
9	181.5	30.0	2414	1 P300_HUMAN	Q09472 homo sapien
10	180.5	29.8	2442	1 CBP_HUMAN	Q92793 homo sapien
11	176	29.1	686	1 BDF1_YEAST	P35817 saccharomyc
12	166.5	27.5	1484	1 CES2_HUMAN	Q9bxf3 homo sapien
13	162	26.8	2056	1 CBP1_CAEEL	P34545 caenorhabdi
14	160.5	26.5	1051	1 TF1A_MOUSE	O64127 mus musculus
15	159.5	26.4	1050	1 TF1A_HUMAN	Q15164 homo sapien
16	158	26.1	837	1 GCL2_HUMAN	Q92930 homo sapien
17	155	25.6	830	1 GCL2_MOUSE	Q9fjd2 mus musculus
18	149	24.6	1972	1 BA2B_HUMAN	Q9uif8 homo sapien
19	148	24.5	2130	1 BA2B_CHICK	Q9del3 gallus gall
20	139.5	23.1	1483	1 BA1B_HUMAN	Q9u1g0 homo sapien
21	136.5	22.6	1556	1 BA1A_HUMAN	Q9nrl2 h bromodoma
22	136	22.5	1872	1 T2D1_HUMAN	P21675 homo sapien
23	134.5	22.2	1479	1 BA1B_MOUSE	Q92777 mus musculus
24	133	22.0	1127	1 TF1G_HUMAN	Q9upn9 homo sapien
25	129	21.3	1850	1 BA2A_MOUSE	Q9lye5 mus musculus
26	129	21.3	1878	1 BA2A_HUMAN	Q9uif9 homo sapien
27	128.5	21.2	2269	1 WDR9_HUMAN	Q9ns16 homo sapien
28	127.5	21.1	2065	1 T2D1_DROME	P51123 drosophila
29	127	21.0	1638	1 BRM_DROME	P25439 drosophila
30	125	20.7	1214	1 BRF3_HUMAN	Q9u1d4 homo sapien
31	123.5	20.4	627	1 BA1A_XENLA	O8uvr5 xenopus lae
32	121.5	20.1	1703	1 SNF2_YEAST	P22082 saccharomyc
33	114.5	18.9	1332	1 SPT7_YEAST	P35177 saccharomyc

```

34 109 18.0 625 1 RSC4_YEAST
35 108 17.9 1058 1 BRD1_HUMAN
36 106 17.5 542 1 BRD1_SCHPO
37 101 16.7 1214 1 BRF1_HUMAN
38 98.5 16.3 1186 1 PKCB_HUMAN
39 95 15.7 733 1 YC40_HUMAN
40 94 15.5 1359 1 STH1_YEAST
41 91.5 15.1 1586 1 SN22_HUMAN
42 89 14.7 1647 1 SN24_HUMAN
43 88 14.5 889 1 RSC2_YEAST
44 86 14.2 928 1 RSC1_YEAST
45 84 13.9 1242 1 TBP7_CAEEL

```

ALIGNMENTS

```

RESULT 1
BRD2_HUMAN
ID BRD2_HUMAN STANDARD; PRT; 801 AA.
AC P25440; O00699; O00700; Q15310; Q969U4;
DT 01-MAY-1992 (Rel. 22, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromodomain-containing protein 2 (RING3 protein) (O27.1.1).
GN BRD2 OR RING3 OR KIAA9001.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RX MEDLINE=92329974; PubMed=1352711;
RA Beck S., Hanson I., Kelly A., Pappin D.J.C., Trowsdale J.;
RT "A homologue of the Drosophila female sterile homeotic (fsh) gene in
RT the class II region of the human MHC.";
RL DNA Seq. 2:203-210(1992).
RN [2]
RP REVISIONS TO N-TERMINUS.
RX MEDLINE=96376536; PubMed=8781126;
RA Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
RT "Phylogeny and structure of the RING3 gene.";
RL Immunogenetics 44:391-396(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
RA Sato S., Nagase T., Seki T., Ishikawa K.-I., Tabata S.;
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Thorpe K.;
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 2 bromodomains.
CC -!- SIMILARITY: Contains 1 ET domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62083; CAA43996.1;
CC EMBL; M80613; AAA68890.1; ALT_INIT.
CC EMBL; P42040; BAA07641.1;
CC EMBL; Z84497; CAC69991.1;
CC EMBL; Z96104; CAC69989.1;
CC EMBL; X96670; CAA65450.1;
CC HSSP; Q92831; 1B91.

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENBL; A3866649; AAL26987.1; -
CC HSP; Q92831; LB91.
CC DR Gene; HGNC:13575; BRD4.
CC DR InterPro; IPR001487; Bromodomain.
CC DR Pfam; PF00439; bromodomain; 2.
CC DR PRINTS; PR00503; BROMODOMAIN.
CC DR SMART; SM00297; BROMO; 2.
CC DR PROSITE; PS00633; BROMODOMAIN 1; 1.
CC DR PROSITE; PS00144; BROMODOMAIN 2; 2.
KW Bromodomain; Repeat; Nuclear protein.
FT DOMAIN 75 147 BROMODOMAIN 1.
FT DOMAIN 368 440 BROMODOMAIN 2.
FT DOMAIN 535 594 LYS-RICH.
FT DOMAIN 692 717 SER-RICH.
FT DOMAIN 703 714 POLY-SER.
FT DOMAIN 738 743 POLY-HIS.
FT DOMAIN 757 761 POLY-PRO.
FT DOMAIN 764 770 POLY-PRO.
FT DOMAIN 771 775 POLY-GLN.
FT DOMAIN 776 783 POLY-PRO.
FT DOMAIN 954 964 POLY-PRO.
FT DOMAIN 974 986 POLY-PRO.
FT DOMAIN 1011 1014 POLY-PRO.
FT DOMAIN 1028 1033 POLY-PRO.
FT DOMAIN 1283 1300 POLY-GLN.
FT DOMAIN 1301 1308 POLY-ALA.
FT DOMAIN 1335 1338 POLY-ARG.
FT CONFLICT 720 721 EM -> GP (IN REF. 2).
SQ SEQUENCE 1362 AA; 152219 MW; D52EFLCF9960907 CRC64;

Query Match 82.6%; Score 500; DB 1; Length 1362;
Best Local Similarity 80.2%; Pred. No. 1.6e-45;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVWKALWKHQPFPQPDVAVKLGPDYHKLKIQPDMGTIKRL 60
Db 55 KPRQNTNQLQYLRLVRLKWKHQPFPQPDVAVKLGPDYHKLKIQPDMGTIKRL 114
QY 61 ENNYWAAACMDQFNTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 115 ENNYWAAACMDQFNTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 165

RESULT 4
FSH DROME STANDARD; PRT; 2038 AA.
AC P13709; P13710;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Female sterile homeotic protein (Fragile-chorion membrane protein).
GS FS(1)H OR FSH.
OS Drosophila melanogaster (Fruit fly).
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC Ephyarodea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=89276730; PubMed=2567251;
RA Haynes S.R., Moser B.A., Bhatia-Dey N., David I.B.;
RT "The Drosophila fish locus, a maternal effect homeotic gene, encodes
RT apparent membrane proteins";
RL Dev. Biol. 134:246-257 (1989).
CC -!- FUNCTION: Required maternally for proper expression of other
CC homeotic genes involved in pattern formation, such as UBX.
CC -!- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.

CC -!- SIMILARITY: Contains 2 bromodomains.
CC -!- SIMILARITY: Contains 1 ET domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M23221; AAA28540.1; -
CC EMBL; M23222; AAA28541.1; ALT_TERM.
CC EMBL; M15762; AAA70424.1; -
CC EMBL; M15763; AAA70423.1; -
CC EMBL; M15764; AAA70422.1; -
CC PIR; A43742; A43742.
CC HSSP; Q92831; LB91.
CC FlyBase; FBgn004656; fs(1)h.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 2.
CC PROSITE; PS00633; BROMODOMAIN 1; 2.
CC PROSITE; PS00144; BROMODOMAIN 2; 2.
KW Developmental protein; Bromodomain; Transmembrane; Repeat.
FT DOMAIN 51 123 BROMODOMAIN 1.
FT DOMAIN 495 567 BROMODOMAIN 2.
FT DOMAIN 945 1106 ET DOMAIN.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 750 770 POTENTIAL.
FT TRANSMEM 790 810 POTENTIAL.
FT TRANSMEM 816 830 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 1731 1751 POTENTIAL.
FT TRANSMEM 1939 1959 POTENTIAL.
FT VARIANT 909 909 G -> A.
FT VARIANT 1022 1022 H -> RXPYY.
SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 82.3%; Score 498; DB 1; Length 2038;
Best Local Similarity 77.5%; Pred. No. 4e-45;
Matches 86; Conservative 13; Mismatches 12; Indels 0; Gaps 0;
QY 1 KPGRTNQLQYLHKVWKALWKHQPFPQPDVAVKLGPDYHKLKIQPDMGTIKRL 60
Db 31 RPRNTNQLQYLKTVWKVWKHQPFPQPDVAVKLGPDYHKLKIQPDMGTIKRL 90
QY 61 ENNYWAAACMDQFNTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 91 ENNYWAAACMDQFNTMTNFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 5
YK82_SCHPO STANDARD; PRT; 727 AA.
ID YK82_SCHPO
AC Q9HGP4; P78808;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical bromodomain protein C631.02.
GN SPAC631.02.
OS Schizosaccharomyces pombe (Fission yeast).
CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.B.,
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 353-713 FROM N.A.
RC STRAIN=PR745;
RX MEDLINE=98162722; PubMed=9501991;
RA Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs";
RL DNA Res. 4:363-369(1997).
CC -!- SIMILARITY: Contains 2 bromodomains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL391713; CAC05484.1; -;
DR EMBL; D89157; BAA13819.1; -;
DR PIR; T42517; T42517.
DR GeneDB_Spombe; SPAC631.02; -;
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS00633; BROMODOMAIN 2; 2.
DR PROSITE; PS00633; BROMODOMAIN 2; 2.
KW Hypothetical protein: Bromodomain; Repeat.
FT DOMAIN 246 318 BROMODOMAIN 1.
FT DOMAIN 408 480 BROMODOMAIN 2.
SQ SEQUENCE 727 AA; 81031 MW; 5CFAT3844CBEEF05 CRC64;
Query Match 33.5%; Score 202.5; DB 1; Length 727;
Best Local Similarity 33.6%; Pred. No. 5.7e-14;
Matches 37; Conservative 29; Mismatches 43; Indels 1; Gaps 1;
QY 2 PGRVTNQLYHKVVKALWQHGFANPPRPQVDVAVKGLPDYHKIIPQMDMTIKERLE 61
DB 228 PPMTEQKHYYH-AMLRQLREGSDIPFPAFVDPVKQNIIDYPTIIXNPDLGLTGMQKFS 286
QY 62 NNYWAASECQDFNTFTNYCYINKPDTDVLVMAQTLKIFLOKVASMP 111
DB 287 SGVYSSAGHFDDNNLMESNCFLYNGTSPGVGKGNLQATFERQLQLP 336
RESULT 6
PCAF_HUMAN

ID PCAF_HUMAN STANDARD; PRT; 832 AA.
AC Q92831;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE F300/CBP-associated factor (EC 2.3.1.-) (P/CAF) (Histone acetylase
PCAF).
GN PCAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, TISSUE SPECIFICITY, AND
RP INTERACTION WITH P300 AND CBP.
RC TISSUE=Liver;
RX MEDLINE=96300317; PubMed=8684459;
RT "A P300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A";
RL Nature 382:319-324(1996).
RN [2]
RP REVISIONS.
RC TISSUE=Liver;
RA Nakatani Y.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RN ENZYMOLOGICAL ACTIVITY.
RX MEDLINE=97100994; PubMed=8945521;
RT Ogrzyzko V., Schiltz R.L., Russanova V., Howard B.H., Nakatani Y.;
RT "The transcriptional coactivators p300 and CBP are histone
RT acetyltransferases";
RL Cell 87:953-959(1996).
RN [4]
RP STRUCTURE BY NMR OF 715-832, AND MUTAGENESIS OF VAL-752; TYR-760;
RP TYR-802 AND TYR-809.
RC TISSUE=Liver;
RX MEDLINE=99292086; PubMed=10365964;
RT Dhalluin C., Carlson J.E., Zeng L., He C., Aggarwal A.K., Zhou M.-M.;
RT "Structure and ligand of a histone acetyltransferase bromodomain";
RL Nature 399:491-496(1999).
CC -!- FUNCTION: Functions as a histone acetyltransferase (HAT) to
CC promote transcriptional activation. Has significant histone
CC acetyltransferase activity with core histones (H3 and H4), and
CC also with nucleosome core particles. Inhibits cell-cycle
CC progression and counteracts the mitogenic activity of the
CC adenoviral oncoprotein E1A.
CC -!- SUBUNIT: Interacts with SIRT1 (By similarity). Interacts with
CC EP300 and CBP.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed but most abundant in
CC heart and skeletal muscle.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U57317; AAC50890.2; -;
DR FDB; IN72; 11-DEC-02.
DR FDB; LCM0; 25-APR-01.
DR FDB; LJM4; 17-JUL-02.
DR Genew; HGNC:8638; PCAF.
DR MIM; 602303; -;
DR GO; GO:0005634; C:nucleus; ISS.
DR GO; GO:0016407; F:acetyltransferase activity; IDA.
DR GO; GO:0004407; F:histone deacetylase activity; TAS.
DR GO; GO:0003712; F:transcription cofactor activity; IPI.
DR GO; GO:0007050; P:cell cycle arrest; TAS.

```

DR GO; GO:0006338; P:chromatin modeling; NAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; IDA.
DR GO; GO:0006473; P:protein amino acid acetylation; TAS.
DR InterPro; IPR001487; GCM5domain.
DR InterPro; IPR000182; GCSacetyltrans.
DR Pfam; PF00583; Acetyltransf; 1.
DR Pfam; PF00439; bromodomain; 1.
DR PRINTS; PR00503; bromodomain.
DR SMART; SM00297; BROMO; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
KW Transcription regulation; Transferase; Nuclear protein;
KW Bromodomain; 3D-structure; Cell cycle.
FT DOMAIN 550 623 ACETYLTRANSFERASE.
FT MUTAGEN 740 810 BROMODOMAIN.
FT MUTAGEN 752 752 V-A: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 760 760 Y-A: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 802 802 Y-S: REDUCED ACETYL-LYSINE BINDING.
FT MUTAGEN 809 809 Y-S: COMPLETE LOSS OF ACETYL-LYSINE BINDING.
FT CONFLICT 716 718 KEK -> SHM (IN REF. 3).
FT CONFLICT 766 766 F -> S (IN REF. 3).
FT CONFLICT 805 805 A -> P (IN REF. 3).
SQ SEQUENCE 832 AA; 92960 MW; 72F516F8BD10CD0C CRC64;

Query Match 30.9%; Score 187; DB 1; Length 832;
Best Local Similarity 36.4%; Pred. No. 3e-12;
Matches 39; Conservative 18; Mismatches 48; Indels 2; Gaps 1;

QY 1 KPGVNTQLQYLVKHYKMKLWKHQFAPRPQPDVAVKGLGPDYHKLIKQPMDMGTIKRL 60
DB 720 KEKRPDQLYSTKLSILQVYKQSAWPFNEPVKRTI--APGYEIVIRPMDLKTMSERL 777
QY 61 ENNYVWAASECMDFNTFTNYCYNNKPTDDIVLMAQTLEKIFLQKV 107
DB 778 KNYVYVKLFMDLQRFVFNCKEYNAASEYKCANILEKPFPSKI 824

RESULT 7
CBP_MOUSE STANDARD; PRT; 2441 AA.
AC P45481;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE CREB-binding protein (EC 2.3.1.48).
GN CREBBP OR CBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94019866; PubMed=8413673;
RA Chivria J.C., Kwok R.P.S., Lamb N., Hagiwara M., Montminy M.R.,
RA Goodman R.H.;
RT "Phosphorylated CREB binds specifically to the nuclear protein CBP.";
RL Nature 365:855-859(1993).
RN [2]
RP INTERACTION WITH NCOA3.
RX MEDLINE=97336097; PubMed=9192892;
RA Torchia J., Rose D.W., Inostroza J., Kamei Y., Westin S., Glass C.K.,
RA Rosenfeld M.G.;
RT "The transcriptional co-activator p/CIP binds CBP and mediates
nuclear-receptor function.";
RL Nature 387:677-684(1997).
CC CC -!- FUNCTION: Acetyltransferase enzyme. Acetylates histones, giving a
specific tag for transcriptional activation. Also acetylates non-
histone proteins, like NCOA3 coactivator. Mediates cAMP-gene
regulation by binding specifically to phosphorylated CREB protein.
CC CBP, as coactivator, augments the activity of phosphorylated CREB
to activate transcription of cAMP-responsive genes (By

```

```

similarity).
-!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
-!- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, PCAF and PML. The
TAZ-type 1 domain interacts with HIF1a. Found in a complex
containing NCOA2; NCOA3; IKKα; IKKβ and IKKγ. Interacts with
NCOA6 coactivator. Probably part of a complex with HIF1a and
EP300 (By similarity).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- SIMILARITY: Contains 1 bromodomain.
-!- SIMILARITY: Contains 1 Z2-type zinc finger.
-!- SIMILARITY: Contains 2 TAZ-type zinc fingers.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
-----
EMBL; S66395; AAE28651.1; -.
PIR; S39161; S39161.
PDB; 1P81; 18-OCT-00.
PDB; 1JJS; 03-OCT-01.
PDB; 1XKH; 06-FEB-02.
PDB; 1KDX; 25-NOV-98.
PDB; 1L8C; 24-APR-02.
TRANSFAC; T01318; -.
MGD; MGI:1098280; Crebbp.
GO; GO:000123; C:histone acetyltransferase complex; IDA.
GO; GO:0005634; C:nucleus; IDA.
GO; GO:0004403; F:histone acetyltransferase activity; IDA.
GO; GO:0005515; F:protein binding; IPI.
GO; GO:0003713; F:transcription co-activator activity; IDA.
GO; GO:0045449; F:regulation of transcription; IDA.
InterPro; IPR001487; Bromodomain.
InterPro; IPR003101; XIX.
InterPro; IPR000197; TAZ finger.
Pfam; PF00439; bromodomain; 1.
Pfam; PF02172; XIX; 1.
Pfam; PF02135; zf-TAZ; 2.
Pfam; PF00569; ZZ; 1.
PRINTS; PR00503; BROMODOMAIN.
SMART; SM00297; BROMO; 1.
SMART; SM00551; ZF-TAZ; 2.
SMART; SM00291; ZNF-ZZ; 1.
PROSITE; PS00633; BROMODOMAIN_1; 1.
PROSITE; PS00014; BROMODOMAIN_2; 1.
PROSITE; PS0134; ZF-TAZ; 2.
PROSITE; PS0135; ZF-ZZ; 1.
PROSITE; PS0135; ZF-ZZ; 2.
Transferase; Transcription regulation; Nuclear protein; Activator;
Bromodomain; Zinc-finger; Repeat; 3D-structure.
ZN_FING 346 432 TAZ-TYPE 1.
DOMAIN 1104 1176 BROMODOMAIN.
ZN_FING 1702 1745 ZZ-TYPE.
ZN_FING 1756 1847 TAZ-TYPE 2.
DOMAIN 1062 1065 POLY-GLU.
DOMAIN 1556 1563 POLY-GLU.
DOMAIN 1944 1949 POLY-PRO.
DOMAIN 1968 1971 POLY-GLN.
DOMAIN 2082 2086 POLY-GLN.
DOMAIN 2200 2216 POLY-GLN.
DOMAIN 2296 2299 POLY-GLN.
SEQUENCE 2441 AA; 265474 MW; 0ABB028C3112F419 CRC64;

Query Match 30.5%; Score 184.5; DB 1; Length 2441;
Best Local Similarity 38.0%; Pred. No. 1.7e-11;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAMWPFQPDVAVKGLGPDYHKLIKQPMDMGTIKRLLENNYVWAASECMQDF 75

```

Db 1099 LEALYRQDPESLPFRQVDPQLGIPDYFDIVKXPMDLSTIKRKLDTGQYQEPNQYVDDV 1158

OY 76 NTMFNFCVYKPTDDIVLMAQTLEKFLQKV 107

Db 1159 RLNFNNAWLNKRTSRVYKFCSCLAEPFEQEI 1190

RESULT 8

GCN5_YEAST

AC Q03330; STANDARD; PRT; 439 AA.

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DE Histone methyltransferase GCN5 (EC 2.3.1.48).

GN GCN5 OR ADA4 OR YOR252W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI TaxID=4932;

RN [1]_SEQUENCE FROM N.A.

RP MEDLINE=93011009; PubMed=1396595;

RX Georgakopoulos T., Thireos G.;

RA "Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription.";

RT EMBO J. 11:4145-4152(1992).

RL [2]

RN [2]_SEQUENCE FROM N.A.

RP MEDLINE=97313265; PubMed=9169869;

RX Tettelin H., Agostoni Carbone M.L., Albermann K., Albers K.,

RA Artzy-Kronin D., Barreiros T., Bertani I., Bjorson A.J.,

RA Bruckner M., Bruschi C.V., Carignani G., Castagnoli L., Cerdan E.,

RA Clemente M.L., Coblenz A., Coglieva M., Coissac E., Defoor E.,

RA Del Bino S., Delius H., Delner D., de Wergifosse P., Dujon B.,

RA Durand P., Entian K.-D., Eraso P., Escribano V., Fabiani L.,

RA Fartmann B., Feroli P., Feuermann M., Frontali L., Garcia-Gonzalez M.,

RA Garcia-Saez M.I., Goffeau A., Guerrero P., Hani J., Hansen M.,

RA Hebling U., Hernandez K., Heumann K., Hilger F., Hofmann B.,

RA Indge K.J., James C.M., Klima R., Koetter P., Kramer B., Kramer W.,

RA Laugin G., Leuther H., Louis E.J., Maillier E., Marconi A.,

RA Martegani E., Mason M.J., Mazzoni C., McReynolds A.D.K.,

RA Melchiorro P., Meves H.-W., Minenkova O., Mueller-Auer S.,

RA Nauracki A., Netter P., Neu R., Nombela C., Oliver S.G., Panzeri L.,

RA Pagnelli B., Plevani P., Portetelle D., Porcillo F., Potier S.,

RA Rodrigues-Pousada C., Rodriguez-Belmonte E., Rodriguez-Torres A.M.,

RA Rose M., Ruzzi M., Saliola M., Sanchez-Perez M., Schaefer B.,

RA Schaefer M., Schaefer M., Schmidheini T., Schreier A., Skala J.,

RA Souciet J.-L., Steensma H.Y., Talla E., Thierry A., Vandenbol M.,

RA van der Aart Q.J.M., Van Dyck L., Vanoni M., Verhasselt P., Voet M.,

RA Volckaert G., Wambutt R., Watson M.D., Weber N., Wedler H.,

RA Wipfli P., Wolf K., Wright L.P., Zaccaria P., Zimmermann M.,

RA Zollner A., Kleine K.;

RT "The nucleotide sequence of Saccharomycetes cerevisiae chromosome VII.";

RL Nature 387:81-84(1997).

RN [3]

RP SEQUENCE OF 1-170 FROM N.A.

PC STRAIN=S288C / FY1679;

RX MEDLINE=9729234; PubMed=9133742;

RA Feroli P., Carignani G., Favanello A., Guerreiro P., Azevedo D.,

RA Rodrigues-Pousada C., Melchiorro P., Panzeri L.,

RA Agostoni Carbone M.L.;

RT "Analysis of a 17.9 kb region from Saccharomycetes cerevisiae chromosome VII reveals the presence of eight open reading frames, including BR1 (TFIIIB70) and GCN5 genes.";

RL Yeast 13:373-377(1997).

RN [4]

RP SEQUENCE OF 170-439 FROM N.A.

PC STRAIN=S288C;

RX MEDLINE=9729233; PubMed=9133741;

RA Mazzoni C., Ruzzi M., Rinaldi T., Solinas F., Montebove F.,

RA Frontali L.;

RT "Sequence analysis of a 10.5 kb DNA fragment from the yeast chromosome VII reveals the presence of three new open reading frames and of a TRNATHR gene.";

RL Yeast 13:369-372(1997).

RN [5]

RP ASSOCIATION WITH ADA2.

RX MEDLINE=95045371; PubMed=7957049;

RA Marcus G.A., Silverman N., Berger S.L., Horiuchi J., Guarente L.;

RT "Functional similarity and physical association between GCN5 and ADA2: putative transcriptional adaptors.";

RL EMBO J. 13:4807-4815(1994).

RN [6]

RP IDENTIFICATION IN A SAGA COMPLEX WITH SPT2; SPT7; SPT8; SPT20; HF11

ADA2; ADA3 AND TRAI.

RX MEDLINE=99102959; PubMed=9885573;

RA Grant P.A., Schieltz D., Pray-Grant M.G., Yates J.R. III,

RA Workman J.L.;

RT "The ATM-related cofactor Tral is a component of the purified SAGA complex.";

RL Mol. Cell 2:863-867(1998).

RN [7]

RP REVIEW.

RX MEDLINE=20400991; PubMed=10940244;

RA Dyda F., Klein D.C., Hickman A.B.;

RT "GCN5-related N-acetyltransferases: a structural overview.";

RL Annu. Rev. Biophys. Biomol. Struct. 29:81-103(2000).

RN [8]

RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 99-262.

RX MEDLINE=99362588; PubMed=10430873;

RA Trivelp R.C., Rojas J.R., Sterner D.E., Venkataramani R.N., Wang L.,

RA Zhou J., Allis C.D., Berger S.L., Marmorstein R.;

RT "Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator.";

RL Proc. Natl. Acad. Sci. U.S.A. 96:8931-8936(1999).

CC -!- FUNCTION: Acetylates Lys-14 of histone H3. Also acetylates Lys-8 and Lys-16 of histone H4 with a lower preference. Acetylation of histones gives a specific tag for epigenetic transcription activation. Operates in concert with certain DNA-binding transcriptional activators such as GCN4 or HAP2/3/4. Acts via the formation of large multiprotein complexes that modify the chromatin.

CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.

CC -!- SUBUNIT: Interacts with ADA2. Part of the ADA/GCN5 complex that consists of HF11/ADA1, ADA2, ADA3, SPT20/ADA5 and GCN5. Component of the SAGA complex, at least composed of SPT2, SPT7, SPT8, SPT20/ADA5, HF11, ADA2, ADA3/NGG1, TRAI and GCN5.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- SIMILARITY: Contains 1 bromodomain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X68628; CAA48602.1; -.

CC EMBL; Z73037; CAA97281.1; -.

CC EMBL; X99228; CAA67614.1; -.

CC PIR; S28051; S28051.

CC PDB; 1YGH; 02-AUG-99.

CC PDB; 156I; 24-NOV-00.

CC GeneOnline; 141564; -.

CC TRANSFAC; T02145; -.

CC SGD; S0003484; GCN5.

CC GO; GO:0000124; C:SAGA complex; IDA.

CC InterPro; IPR001487; Bromodomain.

CC InterPro; IPR00182; GCN5acetyl_trans.

CC Pfam; PF00583; Acetyltransf; 1.

CC Pfam; PF00439; Bromodomain; 1.

CC PRINTS; PR00503; BROMODOMAIN.

CC SMART; SM00297; BROMO; 1.

DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS00014; BROMODOMAIN 2; 1.
KW Transcription regulation; Transferrin; Activator; Chromatin regulator;
KW Trans-acting factor; Nuclear protein; Bromodomain; 3D-structure.
FT ACT_SITE 173 173
FT DOMAIN 344 414
FT STRAND 100 105
FT STRAND 111 127
FT HELIX 129 130
FT TURN 133 141
FT TURN 143 144
FT STRAND 146 152
FT TURN 153 155
FT STRAND 156 166
FT STRAND 167 169
FT TURN 170 170
FT TURN 171 179
FT TURN 181 182
FT TURN 186 187
FT HELIX 188 203
FT STRAND 208 213
FT STRAND 215 217
FT HELIX 218 223
FT TURN 224 225
FT STRAND 227 227
FT STRAND 234 237
FT STRAND 238 238
FT TURN 239 240
FT STRAND 241 241
FT TURN 244 245
FT TURN 248 253
SQ SEQUENCE 439 AA; 51069 MW; 3200730DDCYEF70D CRC64;

Query Match 30.1%; Score 182; DB 1; Length 439;
Best Local Similarity 38.5%; Pred. No. 5.3e-12;
Matches 37; Conservative 16; Mismatches 41; Indels 2; Gaps 1;

QY 16 VMKALWKHQFANFPQVDVAVKGLPDYKIIQPMQMGTKIKRLENNYVWAASECMQDF 75
Db 339 ILTELQNHAAWPFQVW--KEVPDYDFIKPMDLSTMEIKLSNKYQKMFYIDA 396
QY 76 NTMTFCYIYKDDIVLVAQTLKIFLOKVASMP 111
Db 397 RLVENCRWNGENTSYYKIANLEKFFNNKVEIP 432

RESULT 9
P300 HUMAN STANDARD; PRT; 2414 AA.
AC Q09472;
DT 15-JUL-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE EIA-associated protein p300 (EC 2.3.1.48).
GN EP300 OR P300.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011587; PubMed=7523245;
RA Eckner R., Eben M.E., Newsome D., Gerdes M., Decaprio J.A.,
RA Lawrence J.B., Livingston D.M.,
RT "Molecular cloning and functional analysis of the adenovirus EIA-
RT associated 300-KD protein (p300) reveals a protein with properties of
RT a transcriptional adaptor.";
RL Genes Dev. 8:869-884(1994).
RN [2]
RP ENZYMOLOGICAL ACTIVITY.
RX MEDLINE=97100994; PubMed=8945521;
RA Ogryzko V.V., Schiltz R.L., Russanova V., Howard B.H., Nakatani Y.,
RT "The transcriptional coactivators p300 and CBP are histone

RT acetyltransferases";
RL Cell 87:953-959(1996).
RN [3]
RP INTERACTION WITH PCAF.
RX MEDLINE=96300317; PubMed=8684459;
RA Yang X.-J., Ogryzko V.V., Nishikawa J.-I., Howard B.H., Nakatani Y.,
RT "A p300/CBP-associated factor that competes with the adenoviral
RT oncoprotein E1A";
RL Nature 382:319-324(1996).
RN [4]
RP INTERACTION WITH HIF1A AND CREBBP.
RX MEDLINE=97075102; PubMed=8917528;
RA Arany Z., Huang L.E., Eckner R., Bhattacharya S., Jiang C.,
RA Goldberg M.A., Bunn H.F., Livingston D.M.,
RT "An essential role for p300/CBP in the cellular response to hypoxia";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12969-12973(1996).
RN [5]
RP INTERACTIONS WITH NR3C1.
RX MEDLINE=98250578; PubMed=9590696;
RA Fryer C.J., Archer T.K.,
RT "Chromatin remodelling by the glucocorticoid receptor requires the
RT BRG1 complex";
RL Nature 393:88-91(1998).
RN [6]
RP INTERACTION WITH NCOA6.
RX MEDLINE=20283976; PubMed=10823961;
RA Ko L., Cardona G.R., Chin W.W.,
RT "Thyroid hormone receptor-binding protein, an LXXLL motif-containing
RT protein, functions as a general coactivator";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6212-6217(2000).
RN [7]
RP INTERACTION WITH DTX1.
RX MEDLINE=21576166; PubMed=11564735;
RA Yamamoto N., Yamamoto S.-I., Inagaki F., Kawauchi M., Fukamizu A.,
RA Kishi N., Matsuno K., Nakamura K., Weinmaster G., Okano H.,
RA Nakafuku M.,
RT "Role of Deltex-1 as a transcriptional regulator downstream of the
RT Notch receptor";
RL J. Biol. Chem. 276:45031-45040(2001).
RN [8]
RP PHOSPHORYLATION, AND INTERACTION WITH TCF7L2 AND LEF1.
RX MEDLINE=22450614; PubMed=12446687;
RA Hecht A., Stemmler M.P.,
RT "Identification of a promoter-specific transcriptional activation
RT domain at the C-terminus of the Wnt-effector protein TCF4";
RL J. Biol. Chem. 278:3776-3785(2003).
RN [9]
RP VARIANTS BREAST CANCER PRO-827 AND GLY-1013, VARIANT PANCREATIC CANCER
TYR-1650, AND VARIANT COLORECTAL CANCER GLN-2221.
RX MEDLINE=20164332; PubMed=10700188;
RA Gayther S.A., Batley S.J., Linger L., Bannister A., Thorpe K.,
RA Chin S.-F., Daigo Y., Russell P., Wilson A., Soutter H.M.,
RA Delhanty J.D.A., Ponder B.A.J., Kouzarides T., Caldas C.,
RT "Mutations truncating the EP300 acetylase in human cancers";
RL Nat. Genet. 24:300-303(2000).
RN [10]
RP STRUCTURE BY NMR OF 302-418 IN COMPLEX WITH 786-826 OF HIF1A.
RX MEDLINE=21957254; PubMed=11953990;
RA Freedman S.J., Sun Z.-Y., Poy F., Kung A.L., Livingston D.M.,
RA Wagner G., Eck M.J.,
RT "Structural basis for recruitment of CBP/p300 by hypoxia-inducible
RT factor-1 alpha";
RL Proc. Natl. Acad. Sci. U.S.A. 99:5367-5372(2002).
CC -!- FUNCTION: Functions as histone acetyltransferase and regulates
CC transcription via chromatin remodeling. Acetylates all four core
CC histones in nucleosomes. Histone acetylation gives an epigenetic
CC tag for transcriptional activation. Binds to and may be involved
CC in the transforming capacity of the adenovirus E1A protein.
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + histone = CoA + acetyl-histone.
CC -!- SUBUNIT: Interacts with NR3C1, PCAF, NCOA6, LEF1 and TCF7L2. The
CC TAZ-type 1 domain interacts with HIF1A. Probably part of a complex
CC with HIF1A and CREBBP. Interacts with DTX1.
CC -!- SUBCELLULAR LOCATION: Nuclear.

GO; GO:0005634; C:nucleus; IDA.
GO; GO:0030528; F:transcription regulator activity; IMP.
GO; GO:0007151; P:sporulation (sensu Saccharomyces); IMP.
InterPro; IPR001487; Bromodomain.
Pfam; PF00439; bromodomain; 2.
PRINTS; PR00503; BROMODOMAIN.
SMART; SM00297; BROMO; 2.
PROSITE; PS00633; BROMODOMAIN_1; 2.
PROSITE; PS50014; BROMODOMAIN_2; 2.
Transcription regulation; Nuclear protein; Bromodomain; Repeat;
Sporulation.
KW
DOMAIN 165 237 BROMODOMAIN 1.
DOMAIN 332 404 BROMODOMAIN 2.
DOMAIN 522 586 ET DOMAIN.
FT
CONFLICT 8 8 Q -> LC (IN REF. 1).
FT
CONFLICT 93 94 GA -> R (IN REF. 2).
FT
CONFLICT 94 94 A -> P (IN REF. 2).
FT
CONFLICT 282 282 A -> P (IN REF. 1).
FT
CONFLICT 385 385 D -> E (IN REF. 1).
FT
CONFLICT 493 493 A -> R (IN REF. 4).
FT
CONFLICT 586 AA; 76978 MW; 8CCD52F41F91D0DA CRC64;
SEQUENCE 686 AA; 76978 MW; 8CCD52F41F91D0DA CRC64;
Query Match 29.1%; Score 176; DB 1; Length 686;
Best Local Similarity 31.0%; Pred. No. 3.7e-11;
Matches 35; Conservative 25; Mismatches 51; Indels 2; Gaps 38
QY 1 KPGRVNTOLOYLHKVVMKALWKH-QFAMPRPQVDAVKGLGIPYHKIIKQPMDMGTIKR 58
DB 310 KSKRLQQAQKFCQSVKELMAKKAASNYNPFLEVPDPSVNNLPFYFYVKEPMDLGTIAK 369
QY 59 RLENNYYWAAECMODFNTMFTNCYIYNKPDDIVLMAQTLEKIFLOKVASMP 111
DB 370 KLNDAQYQTMEDFERDLRVLFKNCYTFNPDGTIVNMGHRLEEVFNKQADRP 422
RESULT 12
CES2 HUMAN
ID CES2 HUMAN STANDARD; PRT; 1484 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cat eye syndrome critical region protein 2.
GN CECR2 OR KIAA1740.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=21275466; PubMed=11381032;
RA Footz T.K., Brinkman-Mills P., Banting G.S., Maier S.A., Riazzi M.A.,
RA Bridgland L.J., Hu S., Birren B., Minoshima S., Shimizu N., Pan H.,
RA Nguyen T., Fang Y., Fu Y., Ray L., Wu H., Shauli S., Phan S., Yao Z.,
RA Chen F., Huan A., Hu P., Wang Q., Loh P., Qi S., Roe B.A.,
RA McDermid H.E.;
RT "Analysis of the cat eye syndrome critical region in humans and the
RT region of conserved synteny in mice: a search for candidate genes at
RT or near the human chromosome 22 pericentromere.";
RL Genome Res. 11:1053-1070(2001).
RN [2]
RP SEQUENCE OF 346-1484 FROM N.A. (ISOFORM A).
RC TISSUE=Brain;
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:347-355(2000).
RN [3]
RP SEQUENCE OF 168-1484 FROM N.A. (ISOFORM B), AND INTERACTION WITH
RP LRPPRC.
RC TISSUE=Liver;

```

RX MEDLINE=21686162; PubMed=11827455;
RA Liu L., McKeenhan W.L.;
RT "Sequence analysis of LRPPRC and its SEC1 domain interaction partners
RT suggests roles in cytoskeletal organization, vesicular trafficking,
RT nucleocytoplasmic shuttling, and chromosome activity.";
RL Genomics 79:124-136(2002).
CC -!- FUNCTION: May be involved through its interaction with LRPPRC in
CC the integration of cytoskeletal network with vesicular
CC trafficking, nucleocytoplasmic shuttling, transcription, chromosome
CC remodeling and cytokinesis.
CC -!- SUBUNIT: Interacts with LRPPRC.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=A;
CC IsoId=Q9BXF3-1; Sequence=Displayed;
CC Name=B; Synonyms=CECR2B;
CC IsoId=Q9BXF3-2; Sequence=VSP_000571, VSP_000572, VSP_000573;
CC -!- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, thymus,
CC placenta and lung. Expressed at lower level in brain, heart,
CC colon, spleen, kidney.
CC -!- MISCELLANEOUS: Candidate gene for the Cat Eye Syndrome (CES), a
CC developmental disorder associated with the duplication of a 2 Mb
CC region of 22q11.2. Duplication usually takes in the form of a
CC supernumerary bisatellited isodicentric chromosome, resulting in
CC four copies of the region (represents an inv dup(22)(q11)). CES is
CC characterized clinically by the combination of coloboma of the
CC iris and anal atresia with fistula, downslanting palpebral
CC fissures, preauricular tags and/or pits, frequent occurrence of
CC heart and renal malformations, and normal or near-normal mental
CC development.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF336133; AAK15343.1; -.
DR EMBL; AB051527; BAB21831.1; -.
DR EMBL; AF411609; AAL07393.1; -.
DR HSSP; Q92831; I891.
DR Genew; HGNC:1840; CECR2.
DR MIM; 607576; -.
DR GO; GO:000910; P:cytokinesis; NAS.
DR GO; GO:0007010; P:cytoskeleton organization and biogenesis; NAS.
DR GO; GO:0016192; P:vesicle-mediated transport; NAS.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 1.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS0014; BROMODOMAIN 2; 1.
KW Bromodomain; Alternative splicing.
FT DOMAIN 451 521 BROMODOMAIN.
FT DOMAIN 333 337 POLY-GLU.
FT DOMAIN 611 614 POLY-SER.
FT DOMAIN 1250 1253 POLY-PRO.
FT VARSPPLIC 291 318 Missing (in isoform B).
FT FTId=VSP_000571.
FT FTId=VSP_000572.
FT VARSPPLIC 519 526 EXTMSDN -> GKQGRSLC (in isoform B).
FT FTId=VSP_000572.
FT VARSPPLIC 527 1484 Missing (in isoform B).
FT FTId=VSP_000573.
FT CONFLICT 370 389 MISSING (IN REF. 2).
FT CONFLICT 1029 1029 C -> S (IN REF. 2).
FT CONFLICT 1045 1045 R -> W (IN REF. 2).
SQ SEQUENCE 1484 AA; 164214 MW; 049AA844E51AF63F CRC64;
Query Match 27.5%; Score 166.5; DB 1; Length 1484;
Best Local Similarity 37.2%; Pred. No. 8.6e-10;

```

```

Matches 35; Conservative 21; Mismatches 31; Indels 7; Gaps 3;
QY 12 LHKV--VMKALWQHQPWPFPQPDVAVKLGPDYHKIKQPMDMGTIKRLNNYYWAAS 69
DB 443 MYKLVLDVVKA--HKDSWPFLEPVD--ESVAPNYYQIKAPMDISSMEKKGLGYCTKE 497
QY 70 ECMODFTMTFNCYIYNKPTDDIVLMAQTLEKIF 103
DB 498 EFVNDMTKTFNCRKNGESSESYTKNSDNLRCF 531
RESULT 13
CP1L CAEEL STANDARD; PRT; 2056 AA.
ID CBPI CAEEL
AC P34545;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein cbp-1.
GN CBP-1 OR R10B11.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Bristol N2.
RX MEDLINE=94150718; PubMed=79063198;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnenhammer E., Staden R.,
RA Sultun J., Thierley-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Worldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RA Durbin R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=B;
CC IsoId=P34545-1; Sequence=Displayed;
CC Name=A;
CC IsoId=P34545-2; Sequence=VSP_000557;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -!- SIMILARITY: Contains 1 Z2-type zinc finger.
CC -!- SIMILARITY: Contains 2 TAZ-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z29095; CAA82353.2; -.
DR EMBL; Z29095; CAD18875.1; -.
DR PIR; G88564; G88564.
DR WormPep; R10E11.1a; CE28069.
DR WormPep; R10E11.1b; CE21117.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR003101; KIX.
DR InterPro; IPR000197; TAZ_finger.

```

DR InterPro; IPR001965; Znf PHD.
 DR InterPro; IPR000433; Znf_ZZ.
 DR Pfam; PF00439; bromodomain; 1.
 DR Pfam; PF02172; KIX; 1.
 DR Pfam; PF02135; zf-TAZ; 2.
 DR Pfam; PF00569; ZF; 1.
 DR Pfam; PF00503; BROMODOMAIN.
 DR SMART; SM00297; BROMO; 1.
 DR SMART; SM00551; Znf TAZ; 2.
 DR SMART; SM00291; Znf_ZZ; 1.
 DR PROSITE; PS00633; BROMODOMAIN_1; 1.
 DR PROSITE; PS0014; BROMODOMAIN_2; 1.
 DR PROSITE; PS01359; ZF PHD_1; 1.
 DR PROSITE; PS0134; ZF TAZ; 2.
 DR PROSITE; PS01357; ZF_ZZ_1; 1.
 DR PROSITE; PS0135; ZF_ZZ_2; 1.
 KW Bromodomain; Metal-binding; Zinc; Zinc-finger; Alternative splicing;
 KW Repeat.
 FT ZN-FING 399 505 TAZ-TYPE 1.
 FT DOMAIN 881 953 BROMODOMAIN.
 FT ZN-FING 1493 1534 ZZ-TYPE.
 FT ZN-FING 1550 1631 TAZ-TYPE 2.
 FT DOMAIN 1687 2008 GLY/GLN-RICH.
 FT VARSPLIC 467 478 SDTCTTKKCSV -> F (in isoform a).
 FT SEQUENCE 2056 AA; 227179 MW; 949FF4608C634F01 CRC64;
 Query Match 26.8%; Score 162; DB 1; Length 2056;
 Best Local Similarity 35.9%; Pred. No. 3.7e-09;
 Matches 33; Conservative 19; Mismatches 40; Indels 0; Gaps 0;
 QY 16 VMKALWKHQPWPQPDVAVKGLPYHKIKQPMGVIKRLNNYYWASECMQDF 75
 DB 876 VWEKLDKSEDAAPRPVDAKLLNIPDIHEIKRPMLETVHKKLYAGQYQNGQFCDDI 935
 QY 76 NTMTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
 DB 936 WLMLDNLWLNKNSKVYKGLKLEMFVSM 967
 RESULT 14
 TFIIA_MOUSE
 ID TFIIA_MOUSE STANDARD; PRT; 1051 AA.
 AC Q64127; Q64126;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transcription intermediary factor 1-alpha (TFII1-alpha) (Tripartite
 DE motif protein 24).
 GN TFII1 OR TRIM24 OR TFIIA.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Carcinoma;
 RX MEDLINE=95262642; PubMed=7744009;
 RA le Douarin B., Zechel C., Garnier J.-M., Lutz Y., Tora L., Pierrat B.,
 RA Heery D., Gronemeyer H., Chambon P., Lessou R.;
 RT "The N-terminal part of TFII1, a putative mediator of the ligand-
 RT dependent activation function (AF-2) of nuclear receptors, is fused to
 RT B-raf in the oncogenic protein TIF-2";
 RL EMO J. 14:2020-2033(1995).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=2238257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 Scherfeld A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [3]
 RN INTERACTION WITH CBX1 AND CBX3.
 RP MEDLINE=97133299; PubMed=8978696;
 RA le Douarin B., Nielsen A.L., Garnier J.-M., Ichinose H.,
 Jeanmougin F., Lessou R., Chambon P.;
 RT "A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic
 RT control of transcription by nuclear receptors";
 RL EMO J. 15:6701-6715(1996).
 CC -!- FUNCTION: Interacts selectively in vitro with the AP2-activating
 CC domain of the estrogen receptors. Association with DNA-bound
 CC estrogen receptors requires the presence of estradiol (By
 CC similarity).
 CC -!- SUBUNIT: Interacts with NR3C2 (By similarity). Interacts with CBX1
 CC and CBX3.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q64127-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q64127-2; Sequence=VSP_005773;
 CC -!- TISSUE SPECIFICITY: Detected in all adult tissues, with the
 CC highest expression level in testis.
 CC -!- DISPAR: Participates in a chromosomal translocation that produces
 CC a TIF1A-BRAF (T19) oncogene originally isolated from a fufural-
 CC induced hepatoma.
 CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.
 CC -!- SIMILARITY: Contains 1 bromodomain.
 CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 CC EMBL; S78221; AAB34290.1; -;
 CC EMBL; S78219; AAB34289.1; -;
 CC EMBL; BC056959; AAB56959.1; -;
 CC PIR; S55259; S55259.
 CC HSSP; P29590; 1BOR.
 CC TRANSFAC; T02143; -;
 CC TRANSFAC; T02215; -;
 CC MGD; MGI:109275; Trim24;
 CC GO; GO:0005737; Cytoplasm; IDA.
 CC InterPro; IPR003849; Bbox_C.
 CC InterPro; IPR001487; Bromodomain.
 CC InterPro; IPR000315; Znf Bbox.
 CC InterPro; IPR001965; Znf PHD.
 CC InterPro; IPR001841; Znf_ring.
 CC Pfam; PF00439; bromodomain; 1.
 CC Pfam; PF00828; PHD; 1.
 CC Pfam; PF00643; zf-B_Box; 2.
 CC Pfam; PF00097; zf-C3HC4; 1.
 CC PRINTS; PR01406; BBOXZNFINGER.

```

DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00502; BBOX; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS00119; ZF_EBOX; 2.
DR PROSITE; PS01359; ZF_PHD_1; 1.
DR PROSITE; PS00016; ZF_PHD_2; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Transcription regulation; Repressor; DNA-binding; Bromodomain;
KW Zinc-finger; Alternative splicing; Nuclear protein; Coiled coil;
KW Repeat; Proto-oncogene; Chromosomal translocation.
FT DOMAIN 8 15 POLY-ALA.
FT DOMAIN 19 22 POLY-ALA.
FT ZN_FING 52 77 RING-TYPE.
FT ZN_FING 158 211 B_BOX-TYPE 1.
FT ZN_FING 218 259 B_BOX-TYPE 2.
FT DOMAIN 289 359 COILED COIL (POTENTIAL).
FT DOMAIN 344 347 POLY-GLN.
FT DOMAIN 583 587 POLY-SER.
FT DOMAIN 755 780 NUCLEAR RECEPTOR BINDING SITE (NRBS).
FT ZN_FING 827 874 PHD-TYPE.
FT DOMAIN 908 908 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 933 988 BROMODOMAIN.
FT SITE 332 333 BREAKPOINT FOR TRANSLOCATION TO FORM
TIFIA-BRAF ONCOGENE.
FT VARSPLIC 477 510 Missing (in isoform Short).
SQ SEQUENCE 1051 AA; 116656 MW; 610584fC1c6885972 CRC64;
Query Match 26.5%; Score 160.5; DB 1; Length 1051;
Best Local Similarity 38.4%; Pred. No. 2.6e-09;
Matches 33; Conservative 17; Mismatches 31; Indels 5; Gaps 2;
QY 20 LKHWQFAPFROPVDVAVKGLGPDYHKIIKQPMDMGTIKERLENNY--YWAASECQMDFFNT 77
Db 916 LYCHENSFLAFQDP---VPLTVDPYKIKNPMDLSTIKKLQEDICMYTKPDEDFVADFL 972
QY 78 MTNCTNYNKPDTDDIVLMAQTLEKIF 103
Db 973 IFQNCAEFNEPDPSEVANAGIKLESYF 998
RESULT 15
ID TIF1A_HUMAN STANDARD; PRT; 1050 AA.
AC O15164; O93854;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription intermediary factor 1-alpha (TIF1-alpha) (Tripartite
DE motif protein 24).
GN TIF1 OR TRIM24 OR TIFIA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Breast cancer;
RX MEDLINE=97277352; PubMed=9115274;
RA Theriot S., Henriquet C., Rochefort H., Cavailles V.;
RT "Differential interaction of nuclear receptors with the putative human
RL J. Biol. Chem. 272:12062-12068(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=99144725; PubMed=10022127;
RA Venturini L., You J., Stadler M., Gallien R., Lallemand V.,

```

```

RA Koken M.H.M., Mattei M.-G., Ganter A., Chambon P., Losson R.,
RA De The H.;
RT "TIF1gamma, a novel member of the transcriptional intermediary factor
RT 1 family.";
RL Oncogene 18:1209-1217(1999).
RN [3]
RP SEQUENCE OF 477-510 (ISOFORM LONG).
RC TISSUE=Breast cancer;
RA Cavailles V.;
RL Submitted (JAN-1999) to Swiss-Prot.
RN [4]
RP INTERACTION WITH NR3C2.
RX MEDLINE=2038858; PubMed=10935545;
RA Hellal-Levy C., Pagart J., Souque A., Wurtz J.-M., Moras D.,
RA Rafestin-Oblin M.-E.;
RT "Crucial role of the H11-H12 loop in stabilizing the active
RT conformation of the human mineralocorticoid receptor.";
RL Mol. Endocrinol. 14:1210-1221(2000).
CC -!- FUNCTION: Interacts selectively in vitro with the AF2-activating
CC domain of the estrogen receptors. Association with DNA-bound
CC estrogen receptors requires the presence of estradiol.
CC -!- SUBUNIT: Interacts with CBX1 and CBX3 (By similarity). Interacts
CC with NR3C2.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=O15164-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=O15164-2; Sequence=VSP_005772;
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 2 B box-type zinc fingers.
CC -!- SIMILARITY: Contains 1 bromodomain.
CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; AF009353; AAB63585.1; -.
DR EMBL; AF119042; AAD17258.1; -.
DR HSSP; P29590; 1BCR.
DR TRANSPAC; T04945; -.
DR Genew; HGNC:11812; TIF1.
DR MIM; 603406; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0005102; F:receptor binding; TAS.
DR GO; GO:0003713; F:transcription co-activator activity; TAS.
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR003649; Bbox_C.
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR000315; Znf_Box.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00439; bromodomain; 1.
DR Pfam; PF00628; PHD; 1.
DR Pfam; PF00643; zf-B_box; 2.
DR Pfam; PF00097; zf-C3HC4; 1.
DR PRINTS; PR01406; BBOXZNFINGER.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00502; BBOX; 1.
DR SMART; SM00336; BBOX; 2.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00633; BROMODOMAIN_1; FALSE_NEG.
DR PROSITE; PS00014; BROMODOMAIN_2; 1.
DR PROSITE; PS01359; ZF_EBOX; 2.
DR PROSITE; PS01359; ZF_PHD_1; 1.

```

```

DR PROSITE; PS0016; ZF_PHD 2; 1.
DR PROSITE; PS0018; ZF_RING 1; 1.
DR PROSITE; PS0089; ZF_RING 2; 1.
KW Transcription regulation; Repressor; DNA-binding; Bromodomain;
KW Zinc-finger; Alternative splicing; Nuclear protein; Coiled coil;
KW Repeat.
FT DOMAIN 9 15 POLY-ALA.
FT ZN_FING 56 82 RING-TYPE.
FT ZN_FING 158 211 B BOX-TYPE 1.
FT ZN_FING 218 259 B BOX-TYPE 2.
FT DOMAIN 289 359 COILED COIL (POTENTIAL).
FT DOMAIN 344 347 POLY-GLN.
FT DOMAIN 754 779 NUCLEAR RECEPTOR BINDING SITE (NRBS).
FT ZN_FING 826 873 PHD-TYPE.
FT DOMAIN 891 907 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 932 987 BROMODOMAIN.
FT VARSPLIC 477 510 Missing (in isoform Short).
FT CONFLICT 14 20 AASAAAS -> RUGCAP (IN REF. 1).
FT CONFLICT 24 28 SAAPS -> RGG (IN REF. 1).
FT CONFLICT 109 114 GSPVSG -> ARRSA (IN REF. 1).
FT CONFLICT 350 350 A -> T (IN REF. 1).
FT CONFLICT 600 600 D -> N (IN REF. 1).
FT CONFLICT 608 608 M -> I (IN REF. 1).
FT CONFLICT 967 967 A -> R (IN REF. 1).
SQ SEQUENCE 1050 AA; 116831 MW; D341E8022AAC67E CRC64;

Query Match 26.4%; Score 159.5; DB 1; Length 1050;
Best Local Similarity 38.4%; Pred.No. 3.4e-09;
Matches 33; Conservative 17; Mismatches 31; Indels 5; Gaps 2;

QY 20 LWKQFAPFPQPDVAVKGLGPDYHKIKQPMDMGTIKRLENNY--YWLASECQODENT 77
Db 915 LYCHEMSLAFODP--VPLTVPDYIKIKNPMDLUSTIKKRLQEDYSWKPDVFADRL 971

QY 78 MFTNCYIYNKPTDDIVLMAQTLERIF 103
Db 972 IFQCAEFNEPDSEVANAGIKLENYF 997

```

Search completed: July 9, 2004, 13:06:42
Job time : 18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:03:30 ; Search time 39 Seconds
(without alignments)
898.013 Million cell updates/sec

Title: US-09-784-553C-19
Perfect score: 605
Sequence: 1 KGRVTNQYLHKVVMKAL.....IVLMAQTLEKFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTRMBL_25:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	503	11	Q99PC5 mus musculus
2	605	100.0	798	11	O8411 mus musculus
3	605	100.0	798	11	O54795 mus musculus
4	594	98.2	729	13	Q90971 gallus galli
5	561	92.7	539	13	Q72VH4 xenopus lae
6	542	89.6	505	11	O8C665 mus musculus
7	542	89.6	511	11	O8CAX7 mus musculus
8	542	89.6	556	4	Q8NSM3 homo sapien
9	542	89.6	726	11	O8KZF0 mus musculus
10	542	89.6	726	11	Q9J125 mus musculus
11	530	87.6	499	13	Q7SXM5 brachydanio
12	524	86.6	515	13	Q7SXL8 brachydanio
13	521	86.1	701	13	O8OPT7 fugu tubrip
14	517	85.5	558	13	Q7SXP6 brachydanio
15	500	82.6	723	11	O8VHF7 mus musculus
16	500	82.6	731	4	O60433 homo sapien

17	500	82.6	1400	11	Q9ESU6
18	500	82.6	1400	11	O8VHF8
19	500	82.6	1846	4	O86Y88 homo sapien
20	499	82.5	814	13	O8UUM2 deum2 oryzias lat
21	498	82.3	1110	5	O8IRN6 Q9W3L3
22	498	82.3	2038	5	O9W3L3
23	494	81.7	806	13	Q8AWX9
24	491	81.2	732	13	Q9GU61
25	484	80.0	585	13	Q7ZVX9
26	482	79.7	664	5	O8T775
27	459	75.9	463	4	O8IWI6
28	459	75.9	947	4	O14789
29	459	75.9	960	4	O7ZAA6
30	449	74.2	956	11	Q91Y44
31	304	50.2	249	13	Q91686
32	296	48.9	1209	5	O20947
33	281.5	46.5	247	5	O8T326
34	281.5	46.5	513	5	Q9VC06
35	280	46.3	765	5	Q86S79
36	280	46.3	851	5	Q95Y80
37	279	46.1	1087	5	O20948
38	268	44.3	109	13	O57335
39	237	39.2	674	5	Q9V8S1
40	237	39.2	679	5	O8T328
41	221.5	36.6	688	10	Q93YS6
42	221.5	36.6	703	10	Q9LYA2
43	215	35.5	578	3	Q9Y7N0
44	208	34.4	818	10	Q9LK27
45	204	33.7	386	10	Q84XV2

ALIGNMENTS

RESULT 1

Q99PC5 PRELIMINARY; PRT; 503 AA.

AC Q99PC5; 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Ring 3 (Fragment).

GN BRD2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL;

RA	Korf I.;
RT	"Complete sequence of UL26B06.";
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF318183; AAK07919.1; -.
DR	HSSP; Q92831; 1B91.
DR	MGI; MGI:99495; Brd2.
DR	InterPro; IPR001487; Bromodomain.
DR	Pfam; PF00439; bromodomain; 2.
DR	SMART; SM00297; BROMO; 2.
DR	PRINTS; PR00503; BROMODOMAIN.
DR	PROSITE; PS00633; BROMODOMAIN_1; 2.
DR	PROSITE; PS0014; BROMODOMAIN_2; 2.
FT	NON_TER 503
SQ	SEQUENCE 503 AA; 842A566963981F28 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 503;
Best Local Similarity 100.0%; Pred. No. 5e-59;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KGRVTNQYLHKVVMKALWKHQFAPFROPVDAVKLGFLPDYHKIIKQPMDMGTIKRL 60
DB 70 KGRVTNQYLHKVVMKALWKHQFAPFROPVDAVKLGFLPDYHKIIKQPMDMGTIKRL 129

```

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 2
ID C08411 PRELIMINARY; PRT; 798 AA.
AC C08411.
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Female sterile homeotic-related protein Frg-1.
GN BRD2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Rhee K., Brunori M., Besset V., Wolgemuth D.J.;
RT "Expression and potential role of Frg-1, a putative murine
RT bromodomain-containing homologue of the Drosophila gene female sterile
RT homeotic.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF045462; AAC24810.1; -.
DR MGD; MG1:99495; Brd2.
DR Pfam; PF00439; Bromodomain.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00014; BROMODOMAIN_2; 2.
SQ SEQUENCE 798 AA; 88063 MW; A9942517CF15B7A1 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 798;
Best Local Similarity 100.0%; Pred. No. 8.4e-59;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFAPFPQPDVAVKLGDPDYHKIHKQPMDMGTIKRRL 60
DB 70 KPGRTVNLQYLHKVVKALWKHQFAPFPQPDVAVKLGDPDYHKIHKQPMDMGTIKRRL 129

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 3
ID C054795 PRELIMINARY; PRT; 798 AA.
AC C054795;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE RING3 protein.
GN BRD2 OR PSRG1 OR RING3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Thorpe K.L., Beck S.;
RL Immunogenetics 0:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ;
RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,
RA Okumura K., Kimura M., Inoko H.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```

RP SEQUENCE OF 47-549 FROM N.A.
RC STRAIN=INBRED CD-1; TISSUE=Testis;
RA Taniguchi Y., Matsuzaka Y., Fujimoto H., Miyado K., Kohda A.,
RA Okumura K., Kimura M., Inoko H.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=129SVJ;
RA Rowen L., Qin S., Madan A., Loretz C., James R., Dors M., Mix L.,
RA Hall J., Lasky S., Hood L.;
RT "Sequence of the mouse major histocompatibility locus class II
RT region.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL009226; CAA15818.1; -.
DR EMBL; AL009226; CAA15819.1; -.
DR EMBL; D9801; BAA25416.1; -.
DR EMBL; AB010248; BAA24379.1; -.
DR EMBL; AB010247; BAA24378.1; -.
DR EMBL; AB010246; BAA24377.1; -.
DR EMBL; AF100956; AAC69907.1; -.
DR HSSP; Q92831; 1B91.
DR MGD; MG1:99495; Brd2.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00014; BROMODOMAIN_2; 2.
SQ SEQUENCE 798 AA; 88066 MW; 08DD57FBF1385E96 CRC64;

Query Match 100.0%; Score 605; DB 11; Length 798;
Best Local Similarity 100.0%; Pred. No. 8.4e-59;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHQFAPFPQPDVAVKLGDPDYHKIHKQPMDMGTIKRRL 60
DB 70 KPGRTVNLQYLHKVVKALWKHQFAPFPQPDVAVKLGDPDYHKIHKQPMDMGTIKRRL 129

QY 61 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 130 ENNYWAASECMDPNTMFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 180

RESULT 4
QY0971 ID QY0971 PRELIMINARY; PRT; 729 AA.
AC QY0971;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Kinase.
GN RING3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B12;
RX MEDLINE=96376536; PubMed=8781126;
RA Thorpe K.L., Abdulla S., Kaufman J., Trowsdale J., Beck S.;
RT "Phylogeny and structure of the RING3 gene.";
RL Immunogenetics 44:391-396(1996).
DR EMBL; X96669; CAA65449.1; -.
DR HSSP; Q92831; 1B91.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS00014; BROMODOMAIN_2; 2.
SQ SEQUENCE 729 AA; 80389 MW; 06E1B92C804DF7B8 CRC64;

```

```
Query Match      98.2%; Score 594; DB 13; Length 729;
Best Local Similarity 97.3%; Pred. No. 1.3e-57;
Matches 108; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 60
DB 24 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 83
QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 84 ENNYWGAECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 134

RESULT 5
Q7ZYH4 PRELIMINARY; PRT; 539 AA.
AC Q7ZYH4;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to bromodomain-containing 2 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043784; AAH43784.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 1.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 539 539
SQ SEQUENCE 539 AA; 60130 MW; 53841988925415F4 CRC64;

Query Match      92.7%; Score 561; DB 13; Length 539;
Best Local Similarity 89.2%; Pred. No. 4.4e-54;
Matches 99; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 60
DB 71 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 130
QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 131 ENNYWGAECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 181

RESULT 6
Q8C665 PRELIMINARY; PRT; 505 AA.
AC Q8C665;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.

Query Match      89.6%; Score 542; DB 11; Length 511;
Best Local Similarity 87.4%; Pred. No. 5.5e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 60
DB 30 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 89
QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 90 ENNYWGAECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 140
```

```
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK076472; BAC36359.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 505 505
SQ SEQUENCE 505 AA; 55211 MW; 35C76A91AAEC6AB4 CRC64;

Query Match      89.6%; Score 542; DB 11; Length 505;
Best Local Similarity 87.4%; Pred. No. 5.4e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 60
DB 30 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 89
QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 90 ENNYWGAECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 140

RESULT 7
Q8CAX7 PRELIMINARY; PRT; 511 AA.
AC Q8CAX7;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Bromodomain-containing 3 (Fragment).
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK037435; BAC29806.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS50014; BROMODOMAIN_2; 2.
FT NON_TER 511 511
SQ SEQUENCE 511 AA; 55921 MW; E189A3484AEB2679 CRC64;

Query Match      89.6%; Score 542; DB 11; Length 511;
Best Local Similarity 87.4%; Pred. No. 5.5e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 60
DB 30 KPGRTVNLQYLHKVVMKALWKHQFAMFPFQPVDAVKGILGPDYHKIIPQMDMGTIKRL 89
QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 90 ENNYWGAECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 140
```

```

RESULT 8
ID Q8N5M3 PRELIMINARY; PRT; 556 AA.
AC Q8N5M3;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to bromodomain containing 3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032124; AAH32124.1; -.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; Bromodomain; 2.
DR PRINTS; PR00503; Bromodomain.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 556 AA; 60942 MW; 8352F5DF1801A793 CRC64;

Query Match 89.6%; Score 542; DB 4; Length 556;
Best Local Similarity 87.4%; Pred. No. 6.1e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 31 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 90

QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 91 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 141

RESULT 9
ID Q8K2F0 PRELIMINARY; PRT; 726 AA.
AC Q8K2F0;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Bromodomain-containing 3.
GN BRD3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031536; AAH31536.1; -.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; bromodomain; 2.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79791 MW; 4BC2FC0ACA41DECD CRC64;

Query Match 89.6%; Score 542; DB 11; Length 726;
Best Local Similarity 87.4%; Pred. No. 8.2e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 31 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 90

QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 90 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 140

RESULT 10
ID Q9J125 PRELIMINARY; PRT; 726 AA.
AC Q9J125;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Bromodomain-containing FSH-like protein FSRG2.
GN BRD3 OR 241084F24RIK OR FSRG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Shang E., Wolgemuth D.J.;
RL "Cloning and expression pattern of Fsr2, a putative murine bromodomain-containing homolog of the Drosophila gene female sterile homeotic".
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF269193; AAF98072.1; -.
DR HSSP; Q92831; I891.
DR MGD; MGI:1914632; Brd3.
DR InterPro; IPR001487; Bromodomain.
DR Pfam; PF00439; bromodomain; 2.
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BRMO; 2.
DR PROSITE; PS00633; BROMODOMAIN_1; 2.
DR PROSITE; PS0014; BROMODOMAIN_2; 2.
SQ SEQUENCE 726 AA; 79745 MW; 7AB3B4DAD38A78F4 CRC64;

Query Match 89.6%; Score 542; DB 11; Length 726;
Best Local Similarity 87.4%; Pred. No. 8.2e-52;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 60
DB 30 KPGRTNQLQYLVKVMKALWKHGFQFVPPQVDAVKLGLPDYHKIKQPMDMGTIKRL 89

QY 61 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 111
DB 90 ENNYWASACQDFTNFTNCYIYNKPTDDIVLMAQTLKIFLQKVASMP 140

RESULT 11
ID Q7SXMS PRELIMINARY; PRT; 499 AA.
AC Q7SXMS;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

```
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC055533; AAH55533.1; -.
KW Hypothetical protein.
FT NON_TER 499
SQ SEQUENCE 499 AA; 55075 MW; 801BB3D64F9E7216 CRC64;

Query Match 87.6%; Score 530; DB 13; Length 499;
Best Local Similarity 84.7%; Pred. No. 1.2e-50;
Matches 94; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKLIKQPMDMGTIKRRL 60
DB 27 KPGKRTNQLQYMNQVVKTLWKHQFAMFPFYPVDAIKLNPDPYHKVKNPMDMGTIKRRL 86

QY 61 ENNYWAAASECMQDNTMFTNCIYNNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 87 ENNYWTAAGECMQDNTMFTNCIYNNKPTDDIVLMAQALEKIFLOKVAQMP 137

RESULT 12
Q7SXL8 PRELIMINARY; PRT; 515 AA.
AC Q7SXL8;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DE Hypothetical protein (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
```

```
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC055543; AAH55543.1; -.
KW Hypothetical protein.
FT NON_TER 515
SQ SEQUENCE 515 AA; 56253 MW; C3189F56736B32C CRC64;

Query Match 86.6%; Score 524; DB 13; Length 515;
Best Local Similarity 86.5%; Pred. No. 5.7e-50;
Matches 96; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKLIKQPMDMGTIKRRL 60
DB 27 KPGKRTNQLQYMNQVVKTLWKHQFAMFPFYPVDAIKLNPDPYHKIKNPMDMGTIKRRL 86

QY 61 ENNYWAAASECMQDNTMFTNCIYNNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 87 ESNYYASASECMQDNTMFTNCIYNNKPTDDIVLMAQALEKIFLOKVALMP 137

RESULT 13
Q8QFT7 PRELIMINARY; PRT; 701 AA.
AC Q8QFT7;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21845882; PubMed=11856876;
RA Bouchireb N., Grutzner F., Haaf T., Stephens R.J., Elgar G.,
RA Green A.J., Clark M.S.;
RL "Comparative mapping of the human 9q34 region in Fugu rubripes.";
RT Cytoogenet. Cell Genet. 94:173-179(2001).
DR EMBL: AJ311635; CAC84085.1; -.
DR InterPro: IPR001487; Bromodomain.
DR Pfam: PF00439; bromodomain; 2.
DR PRINTS: PR00503; BROMODOMAIN.
DR SMART: SM00297; BROMO; 2.
DR PROSITE: PS00633; BROMODOMAIN_1; 2.
DR PROSITE: PS00014; BROMODOMAIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 701 AA; 76356 MW; 74E1F328615E9BD0 CRC64;

Query Match 86.1%; Score 521; DB 13; Length 701;
Best Local Similarity 84.7%; Pred. No. 1.7e-49;
Matches 94; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 1 KGRVTNQLQYLHKVYKALWKHQFAMFPFQPDVAVKLGDPYHKLIKQPMDMGTIKRRL 60
DB 28 KPGKRTNQLQYMNQVVKTLWKHQFAMFPFYPVDAIKLNPDPYHKIKNPMDMGTIKRRL 87

QY 61 ENNYWAAASECMQDNTMFTNCIYNNKPTDDIVLMAQLEKIFLOKVASMP 111
DB 88 ENNYWAAASECMQDNTMFTNCIYNNKPTDDIVLMAQALEKIFLOKVAQMP 138

RESULT 14
Q7SXP6
```

RESULT 15	Q8VHF7	Q8VHF7	PRELIMINARY;	PRT;	723 AA.
ID	Q8VHF7				
AC	Q8VHF7				
CD	01-MAR-2002	(TREMBLrel. 20, Created)			
DD	01-MAR-2002	(TREMBLrel. 20, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE	Bromodomain-containing protein BRD4 short variant.				
DE	BRD4.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NCBI	taxID=10090;				
[1]	SEQUENCE FROM N.A.				
RP					

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:05:11 ; Search time 23 Seconds
(without alignments)
249.151 Million cell updates/sec

Title: US-09-784-553c-19
Perfect score: 605
Sequence: 1 KPGRTVNLQYLHKVVKAL.....IVLMAQTLEKIFLQKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCUS-COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	754	4	US-09-392-714-20
2	542	89.6	726	4	US-09-392-714-21
3	500	82.6	722	4	US-09-392-714-22
4	459	75.9	947	4	US-09-418-780A-1
5	459	75.9	947	4	US-09-392-714-23
6	369	61.0	65	1	US-08-227-536-5
7	369	61.0	65	5	PCT-US95-04682-5
8	204	33.7	65	1	US-08-227-536-6
9	204	33.7	65	5	PCT-US95-04682-6
10	192	31.7	238	4	US-09-257-179-80
11	184.5	30.5	2441	1	US-08-194-468-2
12	184.5	30.5	2441	3	US-08-361-739-2
13	184.5	30.5	2441	4	US-09-514-247A-8
14	184.5	30.5	2441	4	US-09-686-316-2
15	181.5	30.0	2414	1	US-08-227-536-2
16	181.5	30.0	2414	5	PCT-US95-04682-2
17	180.5	29.8	2442	4	US-09-514-247A-10
18	158	26.1	61	4	US-09-418-710-55
19	153	25.3	61	4	US-09-418-710-42
20	149	24.6	1972	4	US-09-418-710-21
21	142	23.5	59	4	US-09-418-710-53
22	139.5	23.1	1525	4	US-09-418-710-69
23	139.5	23.1	1527	4	US-09-418-710-27
24	139.5	23.1	1531	4	US-09-418-710-29
25	139	23.0	1969	4	US-09-418-710-72
26	136.5	22.6	1673	4	US-09-418-710-70
27	136.5	22.6	1674	4	US-09-418-710-1

28 136 22.5 59 4 US-09-418-710-41 Sequence 41, Appl
29 136 22.5 1872 1 US-08-188-582-14 Sequence 14, Appl
30 136 22.5 1872 1 US-08-646-715-14 Sequence 14, Appl
31 136 22.5 1893 1 US-08-188-582-11 Sequence 11, Appl
32 136 22.5 1893 1 US-08-646-715-11 Sequence 11, Appl
33 130 21.5 1878 4 US-09-418-710-13 Sequence 13, Appl
34 129 21.3 65 1 US-08-227-536-4 Sequence 4, Appl
35 129 21.3 65 5 PCT-US95-04682-4 Sequence 4, Appl
36 129 21.3 196 4 US-09-690-454-94 Sequence 94, Appl
37 129 21.3 1876 4 US-09-418-710-71 Sequence 71, Appl
38 126 20.8 59 4 US-09-418-710-38 Sequence 38, Appl
39 126 20.8 59 4 US-09-418-710-50 Sequence 50, Appl
40 123 20.3 589 3 US-08-942-008-2 Sequence 2, Appl
41 120 19.7 59 4 US-09-418-710-51 Sequence 51, Appl
42 119 19.8 59 4 US-09-418-710-39 Sequence 39, Appl
43 110 18.2 59 4 US-09-418-710-52 Sequence 52, Appl
44 110 18.2 65 1 US-08-227-536-8 Sequence 8, Appl
45 110 18.2 65 5 PCT-US95-04682-8 Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-09-392-714-20
; Sequence 20, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-20

Query Match Similarity 100.0%; Score 605; DB 4; Length 754;
Best Local Similarity 100.0%; Pred. No. 6.5e-71;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTVNLQYLHKVVKALWKHGFAPFPFQPDVAVKLGLDYHKIIKQPMNDGTIKRRL 60
Db 24 KPGRTVNLQYLHKVVKALWKHGFAPFPFQPDVAVKLGLDYHKIIKQPMNDGTIKRRL 83

QY 61 ENNYWAASECQDPNTMTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 84 ENNYWAASECQDPNTMTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 2
US-09-392-714-21
; Sequence 21, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor

; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 726
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-21

Query Match 89.6%; Score 542; DB 4; Length 726;
Best Local Similarity 87.4%; Pred. No. 1.2e-62;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 60
Db 31 KPGRTNQLQYMNQNVVVKLWKHGFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 90

QY 61 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 91 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 141

RESULT 3

US-09-392-714-22
; Sequence 22, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-22

Query Match 82.6%; Score 500; DB 4; Length 722;
Best Local Similarity 80.2%; Pred. No. 4.2e-57;
Matches 89; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 60
Db 55 KPGRTNQLQYLRLVRLKTLKHGFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 114

QY 61 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 115 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 165

RESULT 4

US-09-418-780A-1
; Sequence 1, Application US/09418780A
; Patent No. 6504009
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/09/418,780A

; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-418-780A-1

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.5e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 60
Db 24 KNGRLTNQLQYLQKVLKDLWKHSFSPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 83

QY 61 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 84 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 5

US-09-392-714-23
; Sequence 23, Application US/09392714A
; Patent No. 6686147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-23

Query Match 75.9%; Score 459; DB 4; Length 947;
Best Local Similarity 73.0%; Pred. No. 1.5e-51;
Matches 81; Conservative 16; Mismatches 14; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVKALWKHOFAPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 60
Db 24 KNGRLTNQLQYLQKVLKDLWKHSFSPFPQPDVAVKLGLPDYHKKIKQPMDMGTIKRRL 83

QY 61 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 111
Db 84 ENNYWAASECQDQNTMTFTNCYIYNKPTDDIVLMAQTLEKIFLQKVASMP 134

RESULT 6

US-08-227-536-5
; Sequence 5, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION

;; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
;; NUMBER OF SEQUENCES: 13
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
;; STREET: Ten Post Office Square
;; CITY: Boston
;; STATE: MA
;; COUNTRY: US
;; ZIP: 02109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/227,536
;; FILING DATE: 14-APR-1994
;; CLASSIFICATION: 436
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Williams Ph.D., Kathleen A.
;; REGISTRATION NUMBER: 34,380
;; REFERENCE/DOCKET NUMBER: DFCI-308XX
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 542-2290
;; TELEFAX: (617) 451-0313
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 65 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; PCT-US95-04682-5

Query Match 61.0%; Score 369; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QFAPFPQPVDVAVKGLPVDYHKIIKQPMDMGTIKRLENNYWAASECMQDFTMTFTNCY 83
Db 1 QFAPFPQPVDVAVKGLPVDYHKIIKQPMDMGTIKRLENNYWAASECMQDFTMTFTNCY 60
QY 84 IYNKP 88
Db 61 IYNKP 65

RESULT 7
PCT-US95-04682-5
; Sequence 5, Application PC/TUS9504682
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: NUCLEIC ACID ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04682
; FILING DATE:
; CLASSIFICATION:

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/227,536
;; FILING DATE: 14-APRIL-1994
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Holliday C. Heine, Ph.D.
;; REGISTRATION NUMBER: 34,346
;; REFERENCE/DOCKET NUMBER: DFCI-308XG999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (617) 542-2290
;; TELEFAX: (617) 451-0313
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 65 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; PCT-US95-04682-5
Query Match 61.0%; Score 369; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 QFAPFPQPVDVAVKGLPVDYHKIIKQPMDMGTIKRLENNYWAASECMQDFTMTFTNCY 83
Db 1 QFAPFPQPVDVAVKGLPVDYHKIIKQPMDMGTIKRLENNYWAASECMQDFTMTFTNCY 60
QY 84 IYNKP 88
Db 61 IYNKP 65

RESULT 8
US-08-227-536-6
; Sequence 6, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
; APPLICANT: Eckner, Richard
; APPLICANT: Ewen, Mark
; APPLICANT: Livingston, David
; TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
; TITLE OF INVENTION: FACTOR P300 AND USES OF P300
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
; STREET: Ten Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,536
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Ph.D., Kathleen A.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: DFCI-308XX
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-2290
; TELEFAX: (617) 451-0313
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 65 amino acids
; TYPE: amino acid

us-09-784-553c-19.ra1

Fri Jul 9 16:10:21 2004

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9672
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)-546-4737
; TELEFAX: (619)-546-9392
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2441 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-194-468-2

Query Match          30.5%; Score 184.5; DB 1; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 12
US-08-961-739-2
; Sequence 2, Application US/08961739A
; Patent No. 6063583
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/08/961,739A
; CURRENT FILING DATE: 1997-10-31
; EARLIER APPLICATION NUMBER: US 194,468
; EARLIER FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-08-961-739-2

Query Match          30.5%; Score 184.5; DB 3; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 13
US-09-514-247A-8
; Sequence 8, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko

; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPAL
; FILE REFERENCE: TANIGUCHI=6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: mouse
; US-09-514-247A-8

Query Match          30.5%; Score 184.5; DB 4; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 14
US-09-686-316-2
; Sequence 2, Application US/09686316
; Patent No. 6646115
; GENERAL INFORMATION:
; APPLICANT: Montminy, Marc R.
; TITLE OF INVENTION: Methods for Treating Diabetes Mellitus
; FILE REFERENCE: SALK1650-1
; CURRENT APPLICATION NUMBER: US/09/686,316
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US/08/961,739
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: US 194,468
; PRIOR FILING DATE: 1994-02-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2441
; TYPE: PRT
; ORGANISM: Mus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2441)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-686-316-2

Query Match          30.5%; Score 184.5; DB 4; Length 2441;
Best Local Similarity 38.0%; Pred. No. 7.8e-15;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

QY 17 MKALWKHQ-FAWPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75
Db 1099 LEALYRQDPESLPFRQPDVAVKGLPDYHKIIKQPMDMGTIKRRLNNYYWAASECMQDF 75

QY 76 NTWFTNCYIYNKPTDDIVLMAQTLEKIFLOKV 107
Db 1159 RLMFNNAWLYNKRKTSRVYKFCSKLAEVFEQEI 1190

RESULT 15
US-08-227-536-2
; Sequence 2, Application US/08227536
; Patent No. 5658784
; GENERAL INFORMATION:
```

```

/ APPLICANT: Eckner, Richard
/ APPLICANT: Ewen, Mark
/ APPLICANT: Livingston, David
/ TITLE OF INVENTION: NUCLEIC ACID, ENCODING TRANSCRIPTION
/ TITLE OF INVENTION: FACTOR P300 AND USES OF P300
/ NUMBER OF SEQUENCES: 13
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
/ STREET: Ten Post Office Square
/ CITY: Boston
/ STATE: MA
/ COUNTRY: US
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/227,536
/ FILING DATE: 14-APR-1994
/ CLASSIFICATION: 436
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Williams Ph.D., Kathleen A.
/ REGISTRATION NUMBER: 34,380
/ REFERENCE/DOCKET NUMBER: DFCI-308XX
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 542-2290
/ TELEFAX: (617) 451-0313
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2414 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-227-536-2

```

```

Query Match      30.0%; Score 181.5; DB 1; Length 2414;
Best Local Similarity 38.0%; Pred. No. 1.9e-14;
Matches 35; Conservative 20; Mismatches 36; Indels 1; Gaps 1;

Qy 17 MKALWKHQ-FAMPPRQPDVAVKLGPDYHKIIKQPMDMGTIKRRLENNYYWAAASECMQDF 75
Db 1062 LEALYRQDPESLPFRQPDVPGIIGIPDYFDIVKSPMDLSTIKRKLDGTQYQEPWQYVDDI 1121

Qy 76 NTMETNCYVINKPTDDIVLNAQTLEKIFLQKV 107
Db 1122 WLFNNALWLNKTSRVYKYCSKLSVEFQEI 1153

```

Search completed: July 9, 2004, 13:08:43
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2004, 13:07:41 ; Search time 47 Seconds
(without alignments)
736.658 Million cell updates/sec

Title: US-09-784-553C-19
Perfect score: 605
Sequence: 1 KGRVTNQLQYLHKVVMKAL.....IVLMAQTLEKIFLOKVASMP 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues
Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605	100.0	111	12	US-09-784-553C-19
2	605	100.0	111	15	US-10-209-201C-19
3	605	100.0	801	14	US-10-146-473-42
4	542	89.6	111	12	US-09-784-553C-20
5	542	89.6	111	15	US-10-209-201C-20
6	542	89.6	140	9	US-09-764-864-1572
7	542	89.6	235	9	US-09-764-864-1159
8	500	82.6	1390	12	US-10-092-900A-224
9	498	82.3	111	12	US-09-784-553C-21
10	498	82.3	111	15	US-10-209-201C-21
11	459	75.9	947	14	US-10-293-822-1
12	269	44.5	113	12	US-09-784-553C-23
13	269	44.5	113	15	US-10-209-201C-23
14	256.5	42.4	113	12	US-09-784-553C-25
15	256.5	42.4	113	15	US-10-209-201C-25

16	244	40.3	113	12	US-09-784-553C-24	Sequence 24, Appl
17	244	40.3	113	15	US-10-209-201C-24	Sequence 24, Appl
18	224	37.0	268	12	US-10-424-599-219965	Sequence 219965, A
19	223	36.9	577	12	US-10-425-114-37396	Sequence 37396, A
20	215	35.5	578	15	US-10-369-493-22750	Sequence 22750, A
21	208.5	34.5	128	12	US-10-424-599-244688	Sequence 244688, A
22	204	33.7	386	14	US-10-176-884-13	Sequence 13, Appl
23	204	33.7	386	14	US-10-177-478-4	Sequence 4, Appl
24	199.5	33.0	656	16	US-10-437-963-154093	Sequence 154093, A
25	198	32.7	110	12	US-09-784-553C-7	Sequence 7, Appl
26	198	32.7	110	15	US-10-209-201C-7	Sequence 7, Appl
27	196	32.4	340	12	US-10-425-114-41650	Sequence 41650, A
28	196	32.4	610	12	US-10-424-599-198828	Sequence 198828, A
29	195	32.2	128	9	US-09-764-864-1571	Sequence 1571, Ap
30	195	32.2	167	16	US-10-437-963-116598	Sequence 116598, A
31	195	32.2	791	16	US-10-437-963-172902	Sequence 172902, A
32	192	31.7	237	12	US-10-220-120-243	Sequence 243, App
33	192	31.7	238	9	US-09-729-835-80	Sequence 80, Appl
34	192	31.7	238	16	US-10-373-809-80	Sequence 80, Appl
35	192	31.7	2781	12	US-10-263-929-122	Sequence 122, App
36	190	31.4	273	12	US-10-424-599-226865	Sequence 226865, A
37	190	31.4	346	12	US-10-425-114-64125	Sequence 64125, A
38	189.5	31.3	563	16	US-10-437-963-198539	Sequence 198539, A
39	189	31.2	372	12	US-10-425-114-38433	Sequence 38433, A
40	188	31.1	751	16	US-10-437-963-172901	Sequence 172901, A
41	187	30.9	516	16	US-10-437-963-116164	Sequence 116164, A
42	187	30.9	832	12	US-09-784-553C-2	Sequence 2, Appl
43	187	30.9	832	15	US-10-209-201C-2	Sequence 2, Appl
44	184.5	30.5	112	12	US-09-784-553C-13	Sequence 13, Appl
45	184.5	30.5	112	15	US-10-209-201C-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-09-784-553C-19
; Sequence 19, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784, 553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-553C-19

Query Match	100.0%;	Score 605;	DB 12;	Length 111;
Best Local Similarity	100.0%;	Pred. No. 5.6e-66;		
Matches 111;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	KGRVTNQLQYLHKVVMKALWKHQFAWPPQVDAVKLGLPDYHKIIKQPMDMGTIKRRL	60	
Db	1	KGRVTNQLQYLHKVVMKALWKHQFAWPPQVDAVKLGLPDYHKIIKQPMDMGTIKRRL	60	
Qy	61	ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP	111	
Db	61	ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP	111	
RESULT 2				
US-10-209-201C-19				
; Sequence 19, Application US/10209201C				
; Publication No. US20040009613A1				

RESULT 4
US-09-784-553C-20
; Sequence 20, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-19

Query Match 100.0%; Score 605; DB 15; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.6e-66;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60

QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111

RESULT 3
US-10-146-473-42
; Sequence 42, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-42

Query Match 100.0%; Score 605; DB 14; Length 801;
Best Local Similarity 100.0%; Pred. No. 6.3e-65;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
DB 71 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 130

QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 131 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 181

RESULT 4
US-09-784-553C-20
; Sequence 20, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-553C-20

Query Match 89.6%; Score 542; DB 12; Length 111;
Best Local Similarity 87.4%; Pred. No. 2.9e-58;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60

QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111

RESULT 5
US-10-209-201C-20
; Sequence 20, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-20

Query Match 89.6%; Score 542; DB 15; Length 111;
Best Local Similarity 87.4%; Pred. No. 2.9e-58;
Matches 97; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60
DB 1 KPGRTNQLQYLHKVVMKALWKHQFAPFPQPDVAVKLGPDYHKIHKQPMDMGTIKRRL 60

QY 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111
DB 61 ENNYWAASECQDFTMTFTNCYIYNKPTDDIVLMAQTLEKIFLOKVASMP 111

Query Match 82.6%; Score 500; DB 12; Length 1390;
Best Local Similarity 80.2%; Pred. No. 8.9e-52;
Matches 89: Conservative 10; Mismatches 12; Indels 0

[illegible]

```

RESULT 9
US-09-784-553C-21
; Sequence 21, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-784-553C-21

```

```

RESULT 10
US-10-209-201C-21
; Sequence 21, Application US/10209201C
; Publication No. US2004009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Agarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 111
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-209-201C-21

```

Qy	1	KPGRVNTQIQYILHKVNVKALWKHQFAMWPRQPDVAVKGLPYHKKIKQPMDMGTIKREL	60
Db	1	RGRNTNTQIQYIJKTVKVIWKHHFSPQQPDVADAKKLNLPDYHKIKQPMDMGTIKREL	60
Qy	61	ENNYWAASECQDFNTMTFCVINYKPTDDILVMAQTLEKIFLQVASMPL	111
Db	61	ENNYWASAKETCDFNTMNNCVNWKPGEDVVMVAOTLEKVFLOKISWPL	111

```

RESULT 11
US-10-293-822-1
; Sequence 1, Application US/10293822
; Publication No. US20030083470A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATOR
; FILE REFERENCE: 06501-043001
; CURRENT APPLICATION NUMBER: US/10/293,822
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/418,780
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: PCT/JP98/01782
; PRIOR FILING DATE: 1998-04-17
; PRIOR APPLICATION NUMBER: JAPAN 9/116402
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 947
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-822-1

```

```

RESULT 12
US-09-784-553C-23
; Sequence 23, Application US/09784553C
; Publication NO. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-784-553C-23

```

```
Query Match      44.5%; Score 269; DB 12; Length 113;
Best Local Similarity 46.0%; Pred. No. 7.8e-25;
Matches 52; Conservative 21; Mismatches 38; Indels 2; Gaps 1;
```


Db	1	KKGLSEQKHNGILKLLSKKRAAYAFYKPDVASALGLHDVHDIKHPMDLSTVKR	60
Qy	59	RLNNYWAASECQDMFTMTCYIYNKPTDDIVLMAQTLKFILOKVASMP	111
Db	61	KMENRYDRAQFAADVRLMFNSCYKYNPDHCVYAMAKLODVEFRYKMP	113

```

RESULT 13
US-10-209-201C-23
; Sequence 23, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Aneel
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-209-201C-23

```

[illegible]

```

RESULT 14
US-09-784-553C-25
; Sequence 25, Application US/09784553C
; Publication No. US20040043378A1
; GENERAL INFORMATION:
; APPLICANT: ZHOU, MING-MING
; APPLICANT: AGGARWAL, ANEEL
; TITLE OF INVENTION: METHODS OF IDENTIFYING MODULATORS OF BROMODOMAINS
; FILE REFERENCE: 2459-1-003 CIP
; CURRENT APPLICATION NUMBER: US/09/784,553C
; CURRENT FILING DATE: 2003-07-08
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-784-553C-25

```

```

Query Match      42.4%; Score 256.5; DB 12; Length 113;
Best Local Similarity 48.5%; Pred No. 2.7e-23;
Matches 48; Conservative 20; Mismatches 28; Indels 3; Gaps 2;

Qy 16 VMKALW--KHQ-FAWPQPQPDVAKVLGLPDYHKLIKQPMDMGTIKRLENNYWAASECM 72
      :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

```

Db      15  ILKELFSKHSGYAPPPYKPPDAEVLGLHDYHDI1KPPMDLGTVCXKQDNRKYKAPPEA  71
QY      73  QDFNTNFTNCYINYPETDDIVLMAQTLEKIFLQKVASMP  111
Db      75  ADVRLFTNCYKYNPPDHDVWAMGRKLQDVFEMRYANIP  113

```

```

RESULT 15
US-10-209-201C-25
; Sequence 25, Application US/10209201C
; Publication No. US20040009613A1
; GENERAL INFORMATION:
; APPLICANT: Verdin, Eric
; APPLICANT: Bruland, Joan
; APPLICANT: Ott, Melanie
; APPLICANT: Zhou, Ming-Ming
; APPLICANT: Aggarwal, Anesh
; TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
; FILE REFERENCE: 2459-1-003CIPDIV
; CURRENT APPLICATION NUMBER: US/10/209,201C
; CURRENT FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/784,553
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 09/510,314
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 25
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-209-201C-25

```

	Query Match	42.4%;	Score 256.5;	DB 15;	Length 113;
	Best Local Similarity	48.5%;	Pred No. 2,7e-23;		
	Matches 48;	Conservative 20;	Mismatches 28;	Indels 3;	Gaps 3
Qy	16	VKALM-KHQ-FAPFPQPDVAKLGLPDHYKIIKQPMGNTGTRKRLENNYVWASFCM	72		
Db	15	ILKELEKSHGSYAMPFYKPDVAMGLGHDYHDIILKKPMDLGTVKRMDNREYKSAPEFA	74		
Qy	73	QDNTMTNCYINKPTDDIVLMAQTLKIFLQKVASMP	111		
Db	75	ADVRLFTNCYKYNPDHDDVAMGKQLQDVFEMRYANIP	113		

Search completed: July 9, 2004, 13:13:32
Job time : 48 secs

